

Vysoká škola báňská – Technická univerzita Ostrava

Fakulta elektrotechniky a informatiky

Katedra kybernetiky a biomedicínského inženýrství

**Biometrická identifikace uživatele na základě
rozpoznání a klasifikace obličejových příznaků –
laboratorní úloha**

Biometrical User Identification on a Base of the Facial
Features Recognition and Classification – A Laboratory
Task

Zadání bakalářské práce

Student:

Vojtěch Raška

Studijní program:

B2649 Elektrotechnika

Studijní obor:

3901R039 Biomedicínský technik

Téma:

Biometrická identifikace uživatele na základě rozpoznání a klasifikace
obličejových příznaků – laboratorní úloha
Biometrical User Identification on a Base of the Facial Features
Recognition and Classification – A Laboratory Task

Jazyk vypracování:

čeština

Zásady pro vypracování:

1. Nastudování principů a využitelnosti biometrie pro identifikaci uživatele.
2. Nastudování metod pro extrakci faciálních příznaků, využívaných pro identifikaci obličeje.
3. Rešerše metod, které se využívají pro detekci a rozpoznání obličeje ze statického a dynamického obrazového signálu.
4. Návrh a implementace datové struktury pro ukládání testovacích obrazových záznamů.
5. Návrh a implementace algoritmu pro automatickou detekci obličeje z obrazu.
6. Návrh a implementace algoritmu pro automatické rozpoznání obličeje z obrazu.
7. Testování algoritmů na reálných obrazových datech.
8. Návrh plně automatizovaného SW rozhraní pro biometrickou analýzu obličeje.
9. Návrh laboratorní úlohy pro identifikaci uživatele z obrazových záznamů.
10. Zhodnocení výsledků.

Seznam doporučené odborné literatury:

- [1] VAN METTER, R. L., J. BEUTEL a H. L. KUNDEL.(eds.) *Handbook of Medical Imaging*. vol. 1. Bellingham, Washington: SPIE, 2009. ISBN 978-0819477729.
- [2] DEMIRKAYA, Omer, Musa Hakan ASYALI a Prasanna SAHOO. *Image processing with MATLAB: applications in medicine and biology*. Boca Raton: CRC Press, c2009. ISBN 978-0-8493-9246-7.
- [3] MARQUES, Oge. *Practical image and video processing using MATLAB*. Hoboken, N.J.: Wiley/IEEE Press, c2011. ISBN 978-0470048153.
- [4] EL-BAZ, Ayman S., Xiaoyi JIANG a Jasjit S. SURI. *Biomedical image segmentation: advances and trends*. Boca Raton: CRC Press, 2016. ISBN 978-1482258554.

Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí bakalářské práce: **Ing. Jan Kubíček**

Datum zadání: 01.09.2017

Datum odevzdání: 30.04.2018



doc. Ing. Jiří Koziorek, Ph.D.
vedoucí katedry




prof. Ing. Pavel Brandštetter, CSc.
děkan fakulty

Prohlášení studenta

Prohlašuji, že jsem tuto bakalářskou/diplomovou práci vypracoval samostatně. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

V Ostravě dne: 20. dubna 2018


.....
podpis studenta

Poděkování

Rád bych vyjádřil poděkování vedoucímu mé bakalářské práce, panu Ing. Janu Kubíčkovi, za rady, věcné připomínky a vstřícnost při konzultacích, které mi pomohly k získání potřebných informací a podkladů pro vytvoření této bakalářské práce

Abstrakt:

Tato bakalářská práce pojednává o tématu biometrické identifikace uživatele na základě rozpoznání a klasifikace obličejových příznaků. Práce obsahuje stručný úvod do biometrie, její metody a využití, a popis metod extrakce obličejových příznaků. Dále bakalářská práce obsahuje rešerši recentních metod pro detekci a rozpoznání obličejů. Hlavní náplní práce je návrh a implementace algoritmů pro detekci a rozpoznávání obličejových příznaků. Analyzované algoritmy byly testovány na rozsáhlém datasetu obličejových dat. Součástí analýzy bylo robustní testování v kontextu aditivního šumu, který byl aplikován na obrazová data. V rámci práce bylo navrženo graficko-uživatelské rozhraní, sloužící pro detekci a klasifikaci obličeje. Práce má částečně edukativní charakter, proto součástí práce je sada testovacích úloh pro úlohy analýzy obličejových příznaků z obrazových dat.

Klíčová slova: Detekce obličeje, rozpoznání obličeje, extrakce příznaků, obličejové příznaky,

Abstract:

This bachelor thesis disserts on the theme of Biometrical User Identification on a Base of the Facial Features Recognition and Classification. Thesis includes brief introduction to the biometrics, its methods, usage and extraction of face features. Next part of the bachelor thesis contains research of various existing and new algorithms for face detection and face recognition. Main part of the thesis is to design and implementation of algorithms for facial features detection and recognition. Analysed algorithms were tested on large dataset of facial images. The analysis included robust testing in the context of additive noise applied to image data. Within the thesis a graphical user interface for detecting and classifying the face was made. The work is partly educational, therefore part of the work is a set of test tasks for the analysis of facial features from image data.

Key words: Face detection, face recognition, extraction features, face features

Obsah

Seznam obrázků	8
Seznam tabulek	9
Seznam grafů	9
Úvod	10
1. Teoretická část	11
1.1 Biometrika	11
1.2 Metody extrakce příznaků (obrazových rysů), které se využívají pro identifikaci obličeje. .	11
2. Rešeršní část	13
2.1 Rozpoznání obličeje s nízkým rozlišením pomocí multi-scale blokového CS-LBP a váženého PCA	13
2.2 Odhad ostrosti obličejových snímků pomocí aproximace spektra	16
2.3 Rozpoznání lidské činnosti v reálném čase na základě hloubkových pohybových map	20
2.4 Rozpoznání tváře pomocí Lokálního binárního vzoru (LBP)	23
2.5 Rozpoznání snímků obličeje s nízkým rozlišením pomocí CLPM	26
2.6 Plně automatická normalizace obličeje a jednoduché rozpoznání obličejů bez omezení prostředí	27
3. Praktická část	30
3.1 3 Design systému pro detekci a klasifikaci obličejových příznaků	30
3.2 Zde by se hodila nová podkapitola: Analýza testovacích dat	30
3.3 Návrh algoritmu pro segmentaci a klasifikaci obličejových příznaků	30
3.3.1 Generátory syntetického šumu	31
3.3.2 Detekce obličeje na základě Viola-Jones algoritmu	32
3.3.3 Extrakce faciálních příznaků	34
3.3.4 Tvorba klasifikátoru	35
3.3.5 Návrh Graficko-uživatelského rozhraní	35
3.4 Testování algoritmu	40
3.5 Vyhodnocení výsledků	44
4. Závěr	48
Seznam literatury:	49
Seznam příloh	51

Seznam obrázků

Obrázek 1. Hloubkové pohybové mapy zachycené v čase z různých pohledů	12
Obrázek 2. Statistický histogram založený na mulsti-scale blokový CS-LBP operátoru	13
Obrázek 3. Vývojový diagram navrženého algoritmu	14
Obrázek 4. Některé výsledky rozpoznání tváře z videa	16
Obrázek 5. Příklady ostrých (vlevo) a rozmazaných (vpravo) snímků a jejich logaritmického aplitudového spektra	17
Obrázek 6. Porovnání odhadů různých metod s ostrostití reálných snímků.....	19
Obrázek 7. (a) Závislost kvality klasifikace na ostrostiti 7 (b) Vypočítaný odhad ostrostiti, chybové pruhy označují standardní chybu	20
Obrázek 8. Příklady sekvencí hloubkových map pro golfový švih a výkop dopředu	21
Obrázek 9. Typický systém rozpoznání obličeje	24
Obrázek 10. Jednoduché textury zachycené LBP operátorem	24
Obrázek 11. Původní LBP operátor.....	25
Obrázek 12. (a) Snímek před použitím SSR algoritmu, (b) snímek po použití SSR algoritmu.....	25
Obrázek 13. (a) Míra rozpoznání s rozdílnou velikostí snímku, (b) Výsledky hromadného rozpoznání	26
Obrázek 14. Převedení natočeného obličeje do frontální polohy. (a) Natočený obraz obličeje. (b) Nastavení síťové masky odpovídající obrázku otočeného obličeje. (c) Triangulace do základní (čelní) polohy. (d) Převedený snímek obličeje.	27
Obrázek 15. Lidská hlava z předního pohledu a s natočení při pohledu ze shora	28
Obrázek 16. naše metoda nastavení AMM. (a) Odhadované orientační body. (b) Triangulace počáteční sítě vytvořené časovanými orientačními body. (c) výsledek sestavení. (d) konečný výsledek	29
Obrázek 17. Příklady snímků z upravené databaze Stirling.....	30
Obrázek 18. Blokové schéma navrhovaného algoritmu pro analýzu faciálních příznaků	31
Obrázek 19. Diagram algoritmus pro aplikaci šumu	32
Obrázek 20. Haarovy příznaky příklady, podobnosti Haarových příznaků s obličejem	33
Obrázek 21. Diagram algoritmu pro detekci obličeje	34
Obrázek 22. Diagram algoritmus po extrakci příznaků snímku	34
Obrázek 23. Vzhled graficko-uživatelského rozhraní.....	35
Obrázek 24. Načtení tréninkové databáze	36
Obrázek 25. Zobrazení testovacího snímku.....	37
Obrázek 26. Detekovaný obličej	37
Obrázek 27. Vizualizace extrahovaných HOG vektorů	38
Obrázek 28. Výsledek rozpoznání obličeje	39
Obrázek 29. Volba šumů a parametrů	39
Obrázek 30. Zašuměné snímky: a) Gaussovský šum, b) Šum Poisson, c) Šum Salt & pepper	40
Obrázek 31. Příklady výsledků rozpoznání různých subjektů	41
Obrázek 32. Příklady rozpoznání při aplikaci šumu Poisson a) bez šumu; b) 1x aplikován; c) 2x aplikován; d) 3x aplikován; e) 3x aplikován	42
Obrázek 33. Aplikace šumu salt & pepper a) bez šumu; b) $d = 0,01$; c) $d = 0,15$; d) $d = 0,02$	43
Obrázek 34. Aplikace Gaussovského šumu a) bez šumu; b) $\sigma = 0,00002$; c) $\sigma = 0,00005$; d) $\sigma = 0,0001$	44

Seznam tabulek

Tab. 1 Tři podmnožiny akcí použitých pro datový soubor MSR-action3D	22
Tab. 2 Porovnání testů (%) pro soubor dat MSR-action3D	22
Tab. 3 Míra rozpoznání vertikálních a standardních odchylek (%) naší metody pro datovou sadu MSR-action3D v náhodných testech.....	23
Tab. 4 Průměrná a standardní odchylka času zpracování komponent naší metody.....	23
Tab. 5 Porovnání různých metod testovaných na databázi FERET	29

Seznam grafů

Graf 1. Míra rozpoznání pro různé rozlišení.....	15
Graf 2. Výsledky testování na nativních datech	45
Graf 3. Výsledky testování při aplikaci Gaussovského šumu	46
Graf 4 Výsledky testování při aplikaci šumu Salt & pepper.....	46
Graf 5. Výsledky testování při aplikaci šumu Poisson	47

Úvod

V této bakalářské práci bylo úkolem vytvořit automatizovaný program pro detekci a identifikaci obličeje za pomoci programu MATLAB. Díky stále lepší technologii kamer, umožňující větší rozlišení a vyšší kvalitu snímků při špatném osvětlení, se stává detekce obličeje snažší. Přesto dnešní systémy ještě nejsou dostatečně dokonalé a může je ovlivnit jak již zmíněné špatné osvětlení, tak i úhel natočení obličeje aj. Samozřejmě další problém je výpočetní složitost. Pro identifikaci se využívají obličejové příznaky jako oči, ústa a nos, přesněji porovnání vzdáleností mezi těmito příznaky.

Teoretická část bakalářské práce popisuje stručně pojem biometrika. Jsou zde popsány některé výhody biometrie a znaky které mohou být použity k identifikaci. Následně je teoretická část věnována popsaní základních metod extrakce obličejových rysů užívaných k identifikaci obličeje.

V rešeršní části jsou popsány, které metody detekce a identifikace obličeje na snímcích pořízených rozdílnými způsoby ať už bezpečnostními kamerami nebo pomocí webkamer. Jsou zde popsány výsledky jednotlivých analýz a porovnání s jinými metodami.

Praktická část je věnována návrhu a naprogramování algoritmu pro detekci obličeje a následné identifikaci uživatele ze snímků získaných z volně dostupné databáze. Tyto snímky obsahují různou variaci výrazů obličeje a jeho natočení jak pro muže, tak ženy. Jsou zde popsány metody, použité pro detekci obličeje na snímku a extrakce obličejových příznaků z nativních a zašuměných dat. Dále je praktická část věnována SW rozhraní a popisu jednotlivých funkcí aktivních prvků. Závěrem byl algoritmus otestován na 92 snímcích ze zmíněné databáze a byla zjištěna míra rozpoznání jednotlivých snímků. Výsledky byly následně zaznamenány v této bakalářské práci.

1. Teoretická část

1.1 Biometrika

Jedná se o obor zabývající se identifikace subjektu na základě jeho charakteristických rysů. Dnes se využívá biometrika převážně při autentizaci. Biometrika využívá jedinečných tělesných znaků pro identifikaci osoby. Výhodou tohoto typu autentizace je, že není nutné pamatovat si několika místné kombinace hesel či neustále s sebou nosit snadno zcizitelný token, např. přihlašovací kartu. Biometrická autentizace je rychlou a pohodlnou a velice přesnou metodou, která je navíc levným řešením, vzhledem ke svým neexistujícím pozdějším nákladům. Její hlavní výhodou je skutečnost, že biometrické charakteristické znaky zůstávají během života neměnné a nelze je ukrást či zapomenout. Podstatou všech biometrických systémů je automatizované snímání biometrických charakteristik a jejich následné porovnávání s údaji předem sejmutými. Cílem v oblasti bezpečnosti je vytvoření komplexních systémů založených na kombinaci měření více charakteristik. Tím se bezpečnost těchto systémů mnohonásobně zvýší. Současné biometrické systémy pracují s různými charakteristickými znaky člověka, jako jsou otisk prstu, geometrie tváře, duhovka oka, sítnice oka, geometrie ruky, geometrie prstů, struktura žil na zápěstí, tvar ucha, složky lidského hlasu, lidský pach, DNA, dynamika podpisu a dynamika psaní na klávesnici a další. [1]

1.2 Metody extrakce příznaků (obrazových rysů), které se využívají pro identifikaci obličeje.

Metody extrakce příznaků slouží pro nalezení daných částí obličeje (oči, ústa, vrásky atp.) nalezení těch to částí pak může sloužit k identifikaci osoby nebo emocí. Obličejové příznaky lze rozdělit na trvalé a přechodové. Mezi trvalé příznaky řadíme ty části obličeje, které se na něm vyskytují trvale, tedy oči, ústa, obočí atd. Kdežto přechodové příznaky se na obličeji vyskytují pouze dočasně a vznikají při nějaké deformaci obličeje (např. úsměv).

Mezi základní metody extrakce příznaků patří:

- a) Metoda porovnání pixelů
- b) Gaborovy filtry
- c) Porovnání jasové intenzity
- d) Analýza hlavních komponent (Principal Component Analysis – PCA)
- e) Aktivní model vzhledu (Active Appearance Model – AAM)
- f) Hloubkové pohybové mapy (Deep Motion Map – DMM)

Viz a) Metoda porovnávání pixelů, jak už název napovídá porovnává hodnoty jednotlivých pixelů. Tato metoda je z důvodu velkého množství dat náročná na čas. Přesto je tato metoda poměrně jednoduchá.

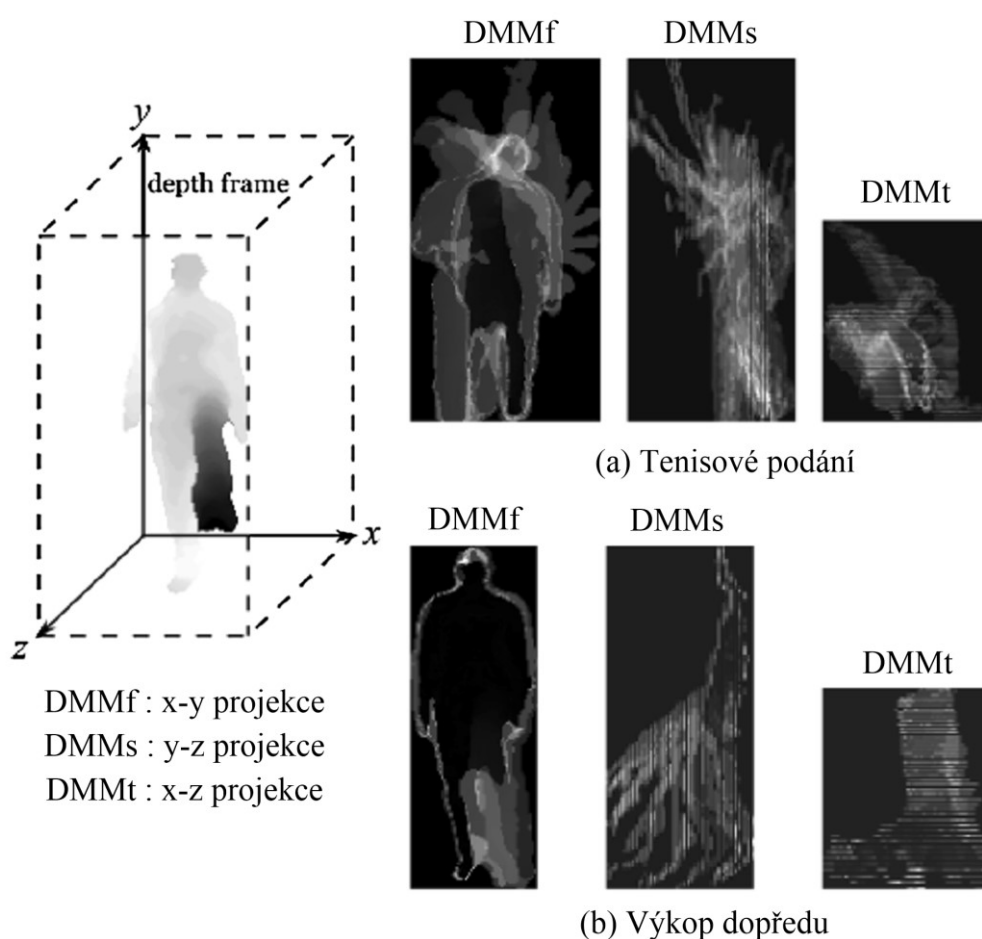
Viz b) Tato metoda funguje podobně jako a) s tím rozdílem, že příznaky neodpovídají hodnotám pixelů, ale hodnotě odezvy obrazových bodů na filtraci Gaborovým filtrem. Výhodou je, že na tuto metodu nepůsobí vlivy osvětlení. Nicméně je tato metoda výpočetně složitější.49[8]

Viz c) Při této metodě dochází k porovnání jasu na hranách tím lze rozlišit jednotlivé části obličeje. Nicméně metoda negativně reaguje na změny světelných podmínek.

Viz d) PCA se používá ke snížení velikosti snímků se zachováním jejích vlastností. Tato metoda projevuje velice dobré výsledky s nízkými výpočetními nároky. Podrobněji je PCA vysvětleno v části 2.1.

Viz e) AAM vychází z aktivních modelů tvaru ASM. ASM jsou vytvořeny pomocí manuálního přidělení bodů na důležité části obličeje. Z těchto ASM se následně vytvořen normalizovaný model tvaru obličeje. Při extrakci se snažíme rozmístit body co nejbližně skutečným pozicím na analyzovaném obličeji. Následně pak dochází k porovnání normalizovaného a analyzovaného obličeje. Tyto fáze se opakují do doby, než dojde ke konvergenci k požadovanému výsledku. Princip můžeme vidět na Obrázek 16. [8]

Viz f) Hluboké pohybové mapy se nepoužívají přímo pro extrakci příznaku, ale pro zaznamenání akce neboli rozdílu pohybu na jednotlivých za sebou jdoucích snímcích



Obrázek 1. Hlubkové pohybové mapy zachycené v čase z různých pohledů

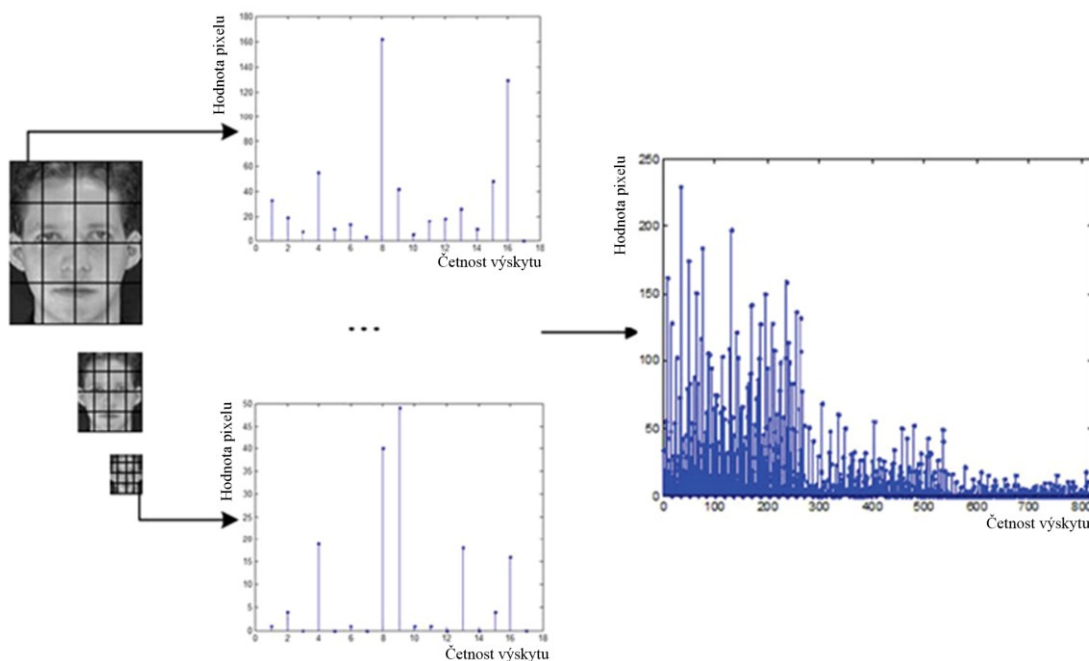
2. Rešeršní část

2.1 Rozpoznání obličeje s nízkým rozlišením pomocí multi-scale blokového CS-LBP a váženého PCA

Tato metoda využívá lokálního binárního vzoru (Local Binary Pattern – LBP) při rozpoznávání tváří s nízkým rozlišením. Přesnější deskriptory a efektivní funkce tváře mohou být extrahovány kombinací multi-scale blokového centra symetrického lokálního binárního vzoru (Center Symmetric Local Binary Pattern – CS-LBP) založeného na Gaussových pyramidách a analýzy hlavní komponenty (Principal Component Analysis – PCA) při nízkém rozlišení.

Metoda LBP byla úspěšně použita pro výpočetně jednoduché a rotační invarianty při popisu a extrakci místní prostorové struktury snímku. Původní verze LBP je tvořena použitím středového pixelu g_c , k redukci pole tvořeného 3x3 pixely na jeden pixel, z důvodu snížení velikosti a zvýšení klasifikace. Kromě toho je původní operátor LBP citlivý na šum obrazu. Proto byl vyvinu operátor CS-LBP, který provádí porovnání mezi sousedními 3 x 3 pixely. Místo porovnání úrovní šedé každého pixelu se středovým pixellem porovnává střed symetrické dvojice pixelů. Rozpoznávání obličeje nemůže dosáhnout vysoké úrovně, pokud použijeme celý obraz pro výpočet histogramu CS-LBP.

Pro získání více informací o obrázcích s nízkým rozlišením, byl navržen model založený na Gaussovských pyramidách. Gaussovy pyramidy jsou sada obrazů získaných opakovaným vyhlazováním a podvzorkováním aktuálních snímků. Lepší reprezentaci snímku obličeje lze získat použitím sady obrazů a vytvořením ztrátové informace o snímku s nízkým rozlišením. Proto vybereme první tři vrstvy pro extrahování funkcí CS-LBP. CS-LBP je vybaven histogramem každého bloku a vrstvy, aby vyjádřily celkovou charakteristiku obrazu. Histogram je znázorněn na Obrázek 2.



Obrázek 2. Statistický histogram založený na mulsti-scale blokový CS-LBP operátoru.

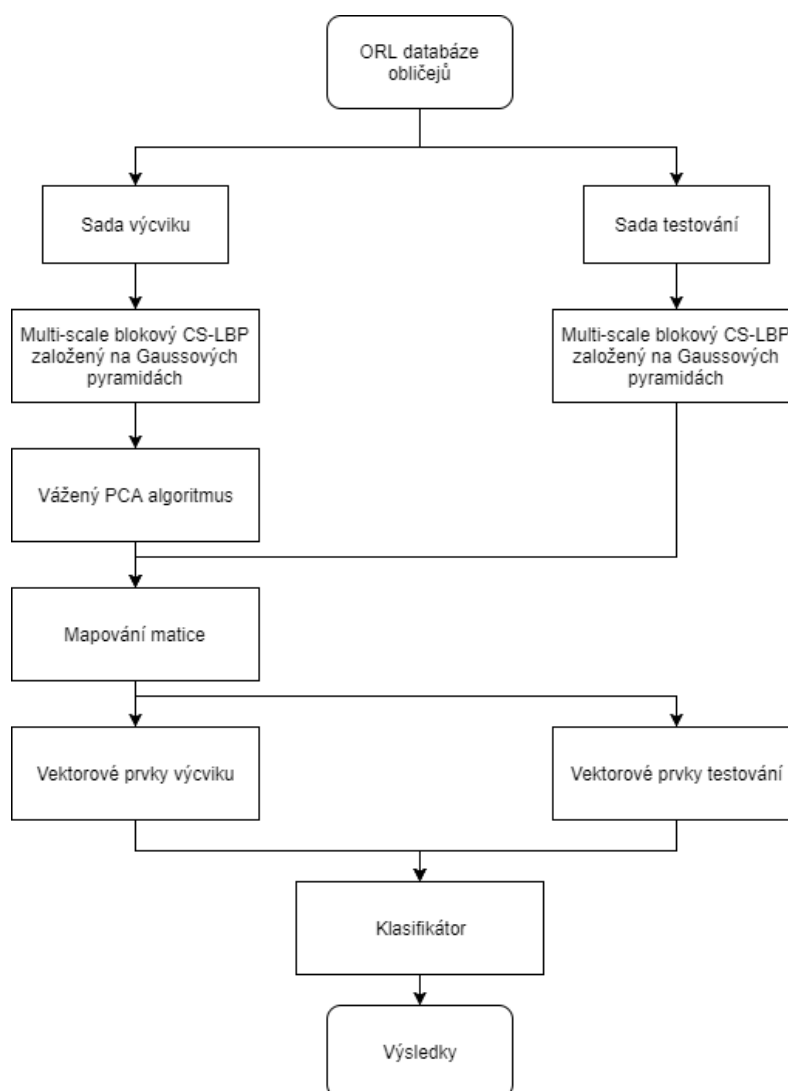
Vážené PCA

Funkce extrahovaná algoritmem blokového CS-LBP obsahovala velké množství redundantních dat. Tento problém způsobil potíže při rozpoznávání a přizpůsobení tváře. Proto je

nutné zmenšit rozměr těchto dat. PCA byla zkoumána pro nízkou výpočetní složitost a vysokou univerzálnost v analýze obrazu a extrakci rysů. PCA lze použít k nalezení podprostoru, jehož kombinací lineárních vektorů mohou být rekonstruovány původní data. V tomto algoritmu, první tři vlastní vektory korespondují s prvními třemi největšími vlastními hodnotami a zdá se, že představují velké změny s různými podmínkami a zhoršují výsledky. Proto algoritmus PCA nemůže v mnoha případech dosáhnout efektivní míry rozpoznání. Za účelem vyřešení výše uvedeného problému je vážený PCA algoritmus navrhován tak aby snížil míru komponent odpovídajících velkým vlastním hodnotám a zvážil míru komponent odpovídajících malým vlastním hodnotám. V případě, že správně zvolíme hodnoty můžeme snížit negativní dopad osvětlení a uchovat užitečné informace o snímcích, pro dosažení lepší míry rozpoznání.

Globální algoritmus

Za účelem zvýšení výpočetní rychlosti a míry rozpoznání CS-LBP metody, stejně jako redukce nepotřebného šumu, byl použitý blokový CS-LBP operátor založený na Gaussových pyramidách a váženém PCA. Navrhovaný proces rozpoznání je znázorněn na Obrázek 3.

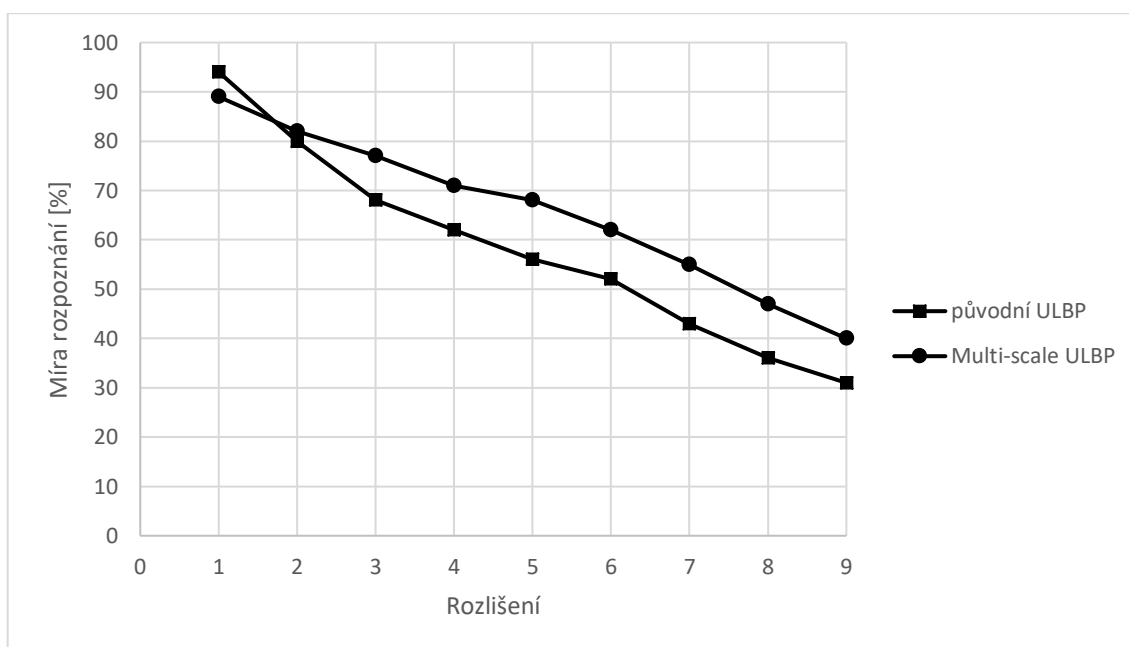


Obrázek 3. Vývojový diagram navrženého algoritmu

Prostřednictvím obecných kroků z Obrázek 3 lze zvolit vhodný blokový parametr a váženou matici. Funkce vypočtené systémem CS-LBP a váženým algoritmem PCA jsou pro identifikaci lepší. Pomocí lokálního operátoru v kombinaci s redukcí globálních rozměrů můžeme dosáhnout uspokojivé míry rozpoznání.

Bylo provedeno několik pokusů s cílem zhodnotit navrhované metody rozpoznání obličeje. Tyto pokusy byly provedeny v databázi obličejů ORL. Tato databáze obsahuje 40 osob s 10 černobílými obličejovými snímky pořízenými pod různým osvětlením, jiným úhlem a výrazem. Rozlišení těchto snímků bylo 112x92 pixelů. Byl použit počítač s 2 GHz procesorem, 4 GB RAM a programem MATLAB.

Databáze je rozdělena do dvou sad, sady výcviku a sady testování. Sada výcviku obsahuje prvních 6 snímků 40 lidí. Zbytek snímků obsahuje testovací sada. Snímky z testovací sady jsou nejprve podvzorkovány na určité rozlišení a následně navzorkovány na původní rozlišení. Gaussovy pyramidy byly získány pomocí sady výcviku a navzorkovaných snímků ze sady testování. Graf 1 zobrazuje výsledek devíti různých rozlišení.



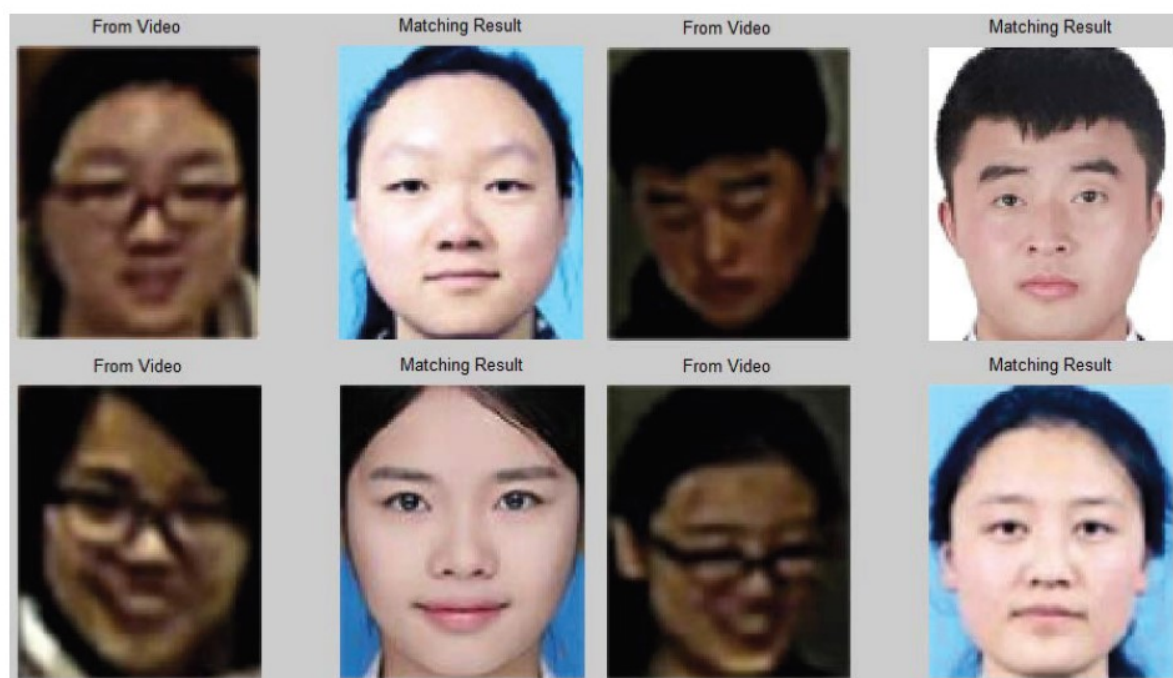
Graf 1. Míra rozpoznání pro různé rozlišení

Jak je vidět na Graf 1, míra rozpoznání tradiční jednotné LBP je výrazně snížena s poklesem rozlišení obrazu. Ačkoli jsou výsledky multi-scale LBP vyšší než originálního LBP u snímků s nižším rozlišením, míra rozpoznání nemůže splnit požadavky praktické aplikace. Proto provádíme experimenty na obrázcích, které byly v databázi ORL třikrát (37x31 px), šestkrát (19x16 px) a devětkrát (12x10 px) zmenšeny.

Výsledky experimentu na video záznamech

Abychom dále ověřili použitelnost navrženého algoritmu, byl proveden experiment s obrazy obličeje z videí. Databáze videí v tomto experimentu byla shromážděna touto skupinou. Testovací sada byla shromážděna pomocí monitorovací kamery a obsahuje 21 snímků se sedmi subjekty. Sada výcviku obsahuje 60 pasových fotek s 20 subjekty s třemi snímky na osobu. Velikost obrazů je

nastavena na 100x120 pixelů. V tomto experimentu vybereme náhodně jeden snímek z testovacího souboru, pro identifikaci nejvíce podobné pasové fotografie a odpovídající jméno ze trénovací sady. Některé výsledky jsou uvedeny na Obrázek 4.



Obrázek 4. Některé výsledky rozpoznání tváře z videa

Výsledky v databázích ORL lidských obličejů ukazují, že míra rozpoznávání může dosáhnout 89,38 % v případě, že rozlišení snímků obličeje klesne na 12x10 pixelů a v podstatě splňuje praktické požadavky na rozpoznávání. Další srovnání dalších deskriptorů a experimentů z videí ukázalo, že nový algoritmus může zlepšit přesnost rozpoznávání.

2.2 Odhad ostrosti obličejových snímků pomocí aproximace spektra

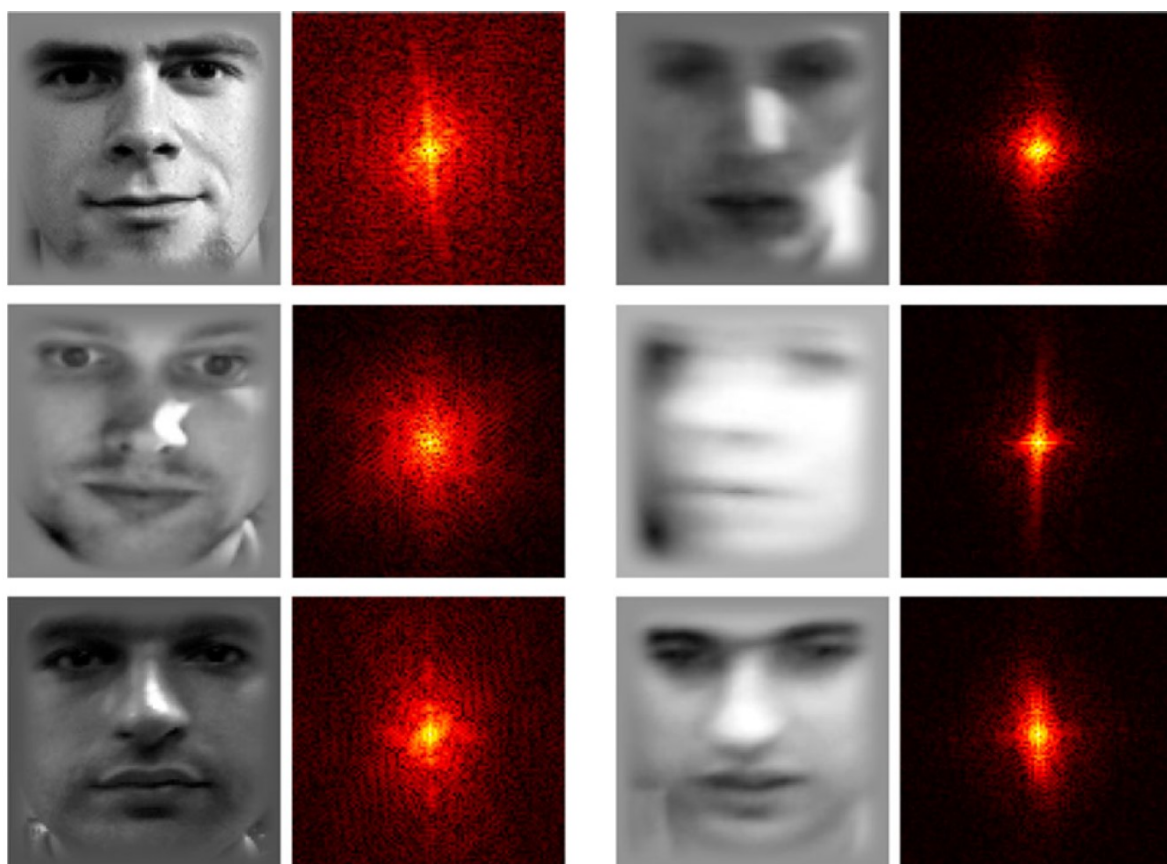
Tento článek představuje nový přístup k hodnocení ostrosti obrazu určený především pro snímky obličeje. Tento přístup lze charakterizovat jako analýzu kmitočtového a amplitudového spektra pomocí přizpůsobení aproximačního modelu a získání odhadu založeného na parametrech modelu.

Aproximace spektra

Mnoho metod odhadu ostrosti využívá skutečnosti, že rozmazání obrazu potlačuje vysoké frekvence spektra. Obrázek 5 ukazuje výsledky uplatnění diskrétní Fourierovy transformace na několik obrazů obličeje, pro porovnání efektu. Obrazy byly zmenšeny na stejnou velikost (100 × 100 pixelů) a na nich byla použita funkce okna. Pro zvětšení hodnot při vyšších frekvencích a normalizaci hodnot do rozsahu [0; 1] se na původní amplitudové spektrum použije následující transformace $S_n(f_x, f_y)$:

$$s_n(f_x, f_y) = \frac{\ln(S_0(f_x, f_y) + 1)}{\ln(S_0(0, 0) + 1)}, \quad (1)$$

kde f_x a f_y jsou prostorové frekvence a $S_n(f_x, f_y)$ je normalizované logaritmické amplitudové spektrum. Pro další analýzu je užitečné reprezentovat spektrum jako řezy v jednotlivých směrech, označené $S_n(f)$ kde $f = \sqrt{f_x^2 + f_y^2}$. Obrázky, které jsou zde prezentovány, lze rozlišit většinou stávajících metod. Některé případy jsou však mnohem náročnější. Můžeme poznamenat, že spektrum zašuměných obrazů se velmi podobné spektru ostrého obrazu a na některých frekvencích jsou zcela nerozlišitelné.



Obrázek 5. Příklady ostrých (vlevo) a rozmazaných (vpravo) snímků a jejich logaritmického amplitudového spektra

Před aproximací byly řezy pro všechny směry spojeny do jednoho, $S_{\text{med}}(f)$, výpočtem střední hodnoty v každém poloměru. Oblíbenou alternativou mediánu v tomto úkolu je aritmetický průměr. Formulování problému aproximace v nejmenším čtverci poskytuje objektivní funkci $J(S * (f))$:

$$J(S^*(f)) = \sum_{i=0}^k ((S_{med}(f_i) - S^*(f_i))^2 \rightarrow \min_{S^*(f)}, \quad (2)$$

Kde $f_0 = 0$ a f_k je maximální frekvence definovaná rozlišením obrazu. K interpretaci vhodné aproximace, bylo použito učení založené na příkladech, které vyžadují soubor obrazů se známou subjektivní ostrostití. Postup, který použitý k získání takových výsledků, byl následující. 15 subjektů bylo požádáno, aby předložili snímky v pořadí vzestupné ostrostiti.

Pro porovnání výsledků této metody byly použity následujícími metody:

- Crété-Roffet et al. [6];
- Marziliano et al. [13];
- the EBM metric by Guan et al. [14];
- De and Masilamani [15];
- the S3 metric by Vu et al. [17].

a dále byly použity některé konvenčně používané metody:

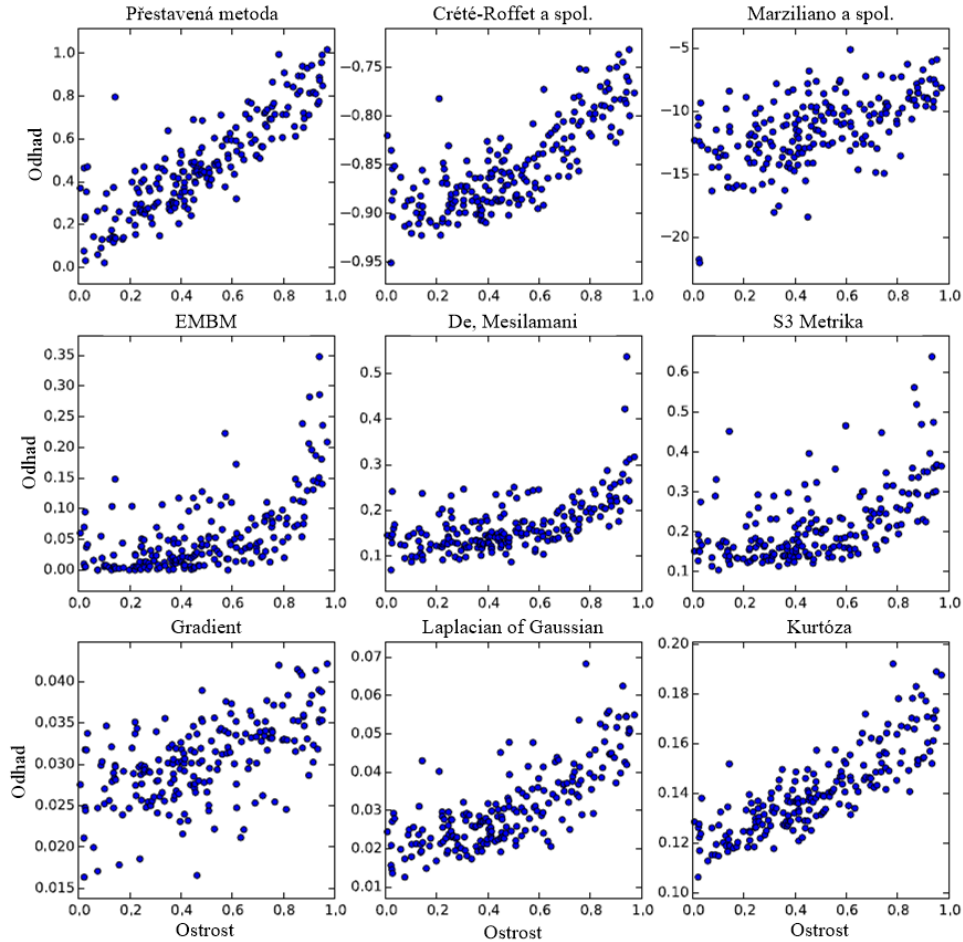
- Součet gradientních veličin ve všech bodech
- Laplacian of Gaussian (LoG): Obraz je nejdříve vyhlazen Gaussovským filtrem, následně je aplikován Laplaceuv operátor, a 0,995 kvantilu výsledků jsou brány jako dostatečná ochrana před šumem.
- Strmost amplitudového spektra. Jedná se o pokles vysokých frekvencí v rozostřených snímcích, které způsobuje zúžení spektra, což obecně odpovídá vyšší strmosti 1D řezu. Je vypočten medián spektra, doplněný jeho zrcadlovou verzí a vypočítáme inverzní strmost tohoto spektra.

Následně pro každý snímek byl získán odhad, použitím modelu Trénovaného na zbývajících vzorcích (varianta metody známá jako cross-validation leave-one-out). Nakonec, hlavní aplikací pro metriku ostrostiti je dobrá / špatná klasifikace, tj. Filtrování rozmazaných obrazů. Pro scénář binární klasifikace je definována metrika výkonu založená na vyvážené přesnosti. Kromě výsledků skóre jednotlivých metod poskytujeme také průměrné výsledky pro srovnání, tj. Průměrné hodnoty skóre pro hodnocení jednotlivých subjektů. Ovšem subjekty, které strávily více času na procesu označování, měly tendenci projevovat vyšší shodu s výsledným skóre, což naznačuje, že dané hodnocení se přibližuje k představě ostrostiti.

V experimentech byly použity 2 datové sady: syntetický datový soubor založený na IMMdatasetu a soubor reálných obrazů. V tomto příspěvku jsou uvedeny pouze výsledky druhého experimentu. Pro oba testy bylo použito stejné nastavení naší metody, jako je počáteční parametr pro vhodný bod, zatímco předpracování a přizpůsobení parametrů pro jiné metody byly provedeny samostatně pro každý soubor dat. Reálná datová sada je tvořena z 200 snímků obličeje pořízených webovou kamerou v různých světelných podmínkách s různými pohyby lidí v prostoru atd. Vzorky těchto obrazů byly ukázány výše na Obrázek 5.

Obrázek 6 ukazuje výsledky zvažovaných metod ve srovnání se subjektivními výsledky. Pro dokonalý odhad bychom očekávali, že všechny body budou ležet na jediné monotónní křivce. V závislosti na aplikaci může nebo nemusí být důležitá, ale obecně je výhodnější univerzální metoda. Z experimentů vyplývá, že navrhovaná metoda překonává existující alternativy jak na umělých, tak na

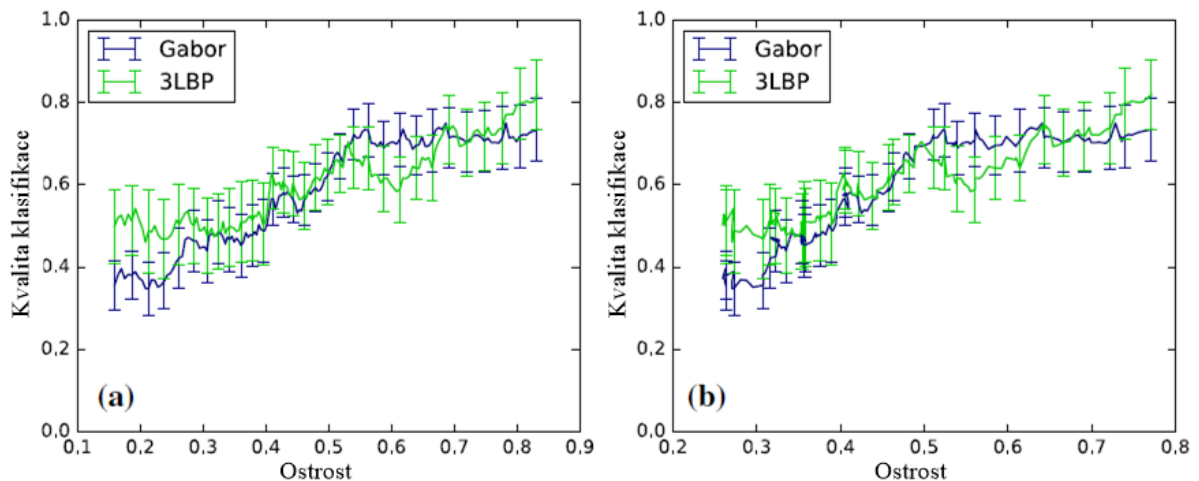
reálných datech, což nám umožňuje dospět k závěru, že model je dostatečně konzistentní se skutečnými deformačními vzory.



Obrázek 6. Porovnání odhadů různých metod s ostrostí reálných snímků

V praktické aplikaci je uvedena analýza pro zlepšení kvality klasifikace obličejů, kterou lze dosáhnout pomocí odhadu ostrosti. Výcviková datová sada obsahuje 500 mužských obličejů a 500 ženských obličejů z datové sady LFW, většinou ostré nebo mírně rozmazané snímky a testovací sada sestávala ze stejných 200 reálných obrazů použitých v předchozí části. Byly použity dva typy deskriptorů: Gaborovy filtry a Místní binární vzory spolu s jejich modifikacemi nazvanými Three-Patch LBP (3LBP) a Four-Patch LBP (4LBP). Závislost střední kvality klasifikace na vnímané ostrosti je znázorněna na Obrázek 7a, b a zobrazuje analogický graf pro navrhovanou míru ostrosti. V obou případech dochází k nárůstu kvality s rostoucí ostrostí, což nám umožňuje zvýšení klasifikace.

Experimenty ukázaly, že výsledky metod v souladu se subjektivní ostrostí jsou lepší než výsledky jiných existujících metod, pokud jde o tři různé cíle: hodnocení obrazu, odhad skutečného skóre ostrosti a klasifikaci dobrých či špatných obrazů. Toto pozorování platí jak pro syntetické testy s umělým rozmazáním a šumem, tak pro test v reálném světě. Také jsme dosáhli zlepšení klasifikace tím, že program vyloučil rozostřené obrazy s nízkým hodnocením.



Obrázek 7. (a) Závislost kvality klasifikace na ostrosti 7 (b) Vypočítaný odhad ostrosti, chybové pruhy označují standardní chybu

2.3 Rozpoznání lidské činnosti v reálném čase na základě hloubkových pohybových map

V tomto příspěvku je zkoumán problém rozpoznání lidských akcí ze sekvencí hloubkových map z hlediska výpočetní efektivity. Tyto snímky jsou zachyceny kamerou RGBD. Konkrétně jsou jako deskriptory vlastností použity hloubkové mapy (DMM) generované hromaděním pohybové energie promítaných hloubkových map ve třech projekčních pohledech (čelní, boční pohled a pohled shora). Ve srovnání s 3D hloubkovými mapami, jsou DMM 2D obrazy, které poskytují kódování pohybových charakteristik akce. Ve srovnání s konvenčními obrazy jsou hloubkové mapy necitlivé na změny světelných podmínek a mohou poskytnout 3D informace k odlišným činnostem, které je obtížné charakterizovat pomocí konvenčních obrazů. Obrázek 8 ukazuje dva příklady skládající se z devíti hloubkových map při golfovém švih (a) Golf swing) a při výkopu dopředu ((b) Forward kick).



(a) *Golf swing*



(b) *Forward kick*

Obrázek 8. Příklady sekvencí hloubkových map pro golfový švih a výkop dopředu

Pomocí hloubkové mapy je možné zachytit informace o struktuře a tvaru 3D objektů. Projekce probíhá ve třech kartézských ortogonálních rovinách. Tento způsob je výpočetně jednoduchý nicméně pro získání DMM je nutno jej upravit. Přesněji řečeno, každý 3D hloubkový rámec se používá k vygenerování tří 2D map odpovídajících přednímu, bočnímu a hornímu pohledu. Pro každou projekční mapu každého směru je vypočítána pohybová energie jako absolutní hodnota ze dvou po sobě jdoucích map. Pro sekvenci hloubky videa s N rámci se DMM_v získá složením energie pohybu z celé videosekvence. Je třeba poznamenat, že pro generování DMM nejsou použity všechny rámce ve videosekvenci. Velikosti jednotlivých pohledu jsou dány velikostí matic $m \times n$, což značí šířku a výšku jednotlivých rámců v pixelech. Vzhledem k tomu, že se hodnoty pixelů používají k dalším výpočtům, byly normalizovány mezi hodnoty 0 a 1. Tím se eliminovaly velké hodnoty pixelů pro další výpočty. Z těchto hodnot jsou později vypočteny vektory jednotlivých DMM. A z těchto vektorů se složí konečný vektor celé DMM. Výsledkem je 4D vektor celé akce. Na závěr je provedena změna velikosti celé DMM. Tímto procesem se značně sníží výpočetní složitost procesu extrakce znaků.

Tato metoda byla použita na datovou sadu Microsoft Research (MSR) Action3D se sekvencí hloubky map zachycenými RGBD kamerou.

Datová sada obsahuje 20 lidských akcí předváděných 10 subjekty. Každý subjekt provedl každou akci dvakrát až třikrát a vždy trochu jinak (změna rychlosti pohybu atd.). Rozlišení hloubkových map bylo 320 x 240 pixelů. Akce byly rozděleny do tří podskupin uvedených v Tab. 1. Pro každou podmnožinu akce byly provedeny tři různé testy.

Akce byly rozděleny do tří podskupin uvedených v Tab. 1 Pro každou podmnožinu byly provedeny tři různé testy. V prvním testu byla použita 1/3 vzorku k výcviku a zbytek vzorku k testování. V druhém testu byly použity 2/3 vzorků k výcviku a zbytek k testování v křížovém testu (třetí test) byla polovina vzorků použita k výcviku a druhá polovina k testování.

Tab. 1 Tři podmnožiny akcí použitých pro datový soubor MSR-action3D

Sada akcí 1 (AS1)	Sada akcí 2 (AS2)	Sada akcí 3 (AS3)
Horizontální vlna (2)	Vysoká vlna (1)	Hod do výšky (6)
Kladivo (3)	Chytání míčku (4)	Výkop dopředu (14)
Úder dopředu (5)	Kreslení X (7)	Kop do boku (15)
Hod do výšky (6)	Kreslení V (8)	Jogging (16)
Tlesknutí (10)	Kreslení kruhu (9)	Tenisový švih (17)
Předklon (13)	Mávání rukou (11)	Tenisové podání (18)
Tenisové podání (18)	Výkop dopředu (14)	Golfový úder (19)
Chycení a hod míčku (20)	Uder ze strany (12)	Chycení a hod míčku (20)

Pro každou sekvenci videa bylo odebráno prvních pět a posledních pět snímků ze dvou důvodů. Za prvé, na začátku a na konci většinou subjekty stály a vykonávaly pouze malý tělesný pohyb. Za druhé, při zpracování docházelo k zvýraznění stojící postavy a tlumení pohybových artefaktů, což přispělo k velkému množství chyb při rekonstrukci.

Pro sadu výcviku a testování byla použita analýza hlavní komponenty (principal component analysis – PCA) pro snížení velikosti. Transformační matice PCA byla vypočtena s využitím sady výcviku a poté použita na sadu testování.

Byla přidána volba, která odmítla danou akci, která nepatřila k sadě akcí. Například, protože akce Skok nebyla zahrnuta do souboru dat MSR-action3D, tak byla zamítnuta. K nalezení vhodného prahu odmítnutí byly provedeny náhodné testy na datové sadě MSR-action3D opakováním každého testu pro každou podmnožinu 200krát.

Tato metoda byla porovnána s existujícími metodami s použitím datové sady MSR-Action3D. Výsledky porovnání jsou uvedeny v Tab. 2. Z Tab. 2 jde vidět, že naše metoda překonala metodu Li et al. ve všech testech. Nejlepší dosažená míra rozpoznání v Tab. 2 je zvýrazněna tučně.

Tab. 2 Porovnání testů (%) pro soubor dat MSR-action3D

	Li a kol.	Lu a kol.	Yang a kol.	Yang a kol.	Vieira a kol.	Naše metoda
První test						
AS1	89,5	98,5	94,7	97,3	98,2	97,3
AS2	89,0	96,7	95,4	92,2	94,8	96,1
AS3	96,3	93,5	97,3	98,0	97,4	98,7
Průměr	91,6	96,2	95,8	95,8	96,8	97,4
Druhý test						
AS1	93,4	98,6	97,3	98,7	99,1	98,6
AS2	92,9	97,2	98,7	94,7	97,0	98,7

AS3	96,3	94,9	97,3	98,7	98,7	100
Průměr	94,2	97,2	97,8	97,4	98,3	99,1
Křížový test						
AS1	72,9	88,0	74,5	96,2	84,7	96,2
AS2	71,9	85,5	76,1	84,1	81,3	83,2
AS3	79,2	63,6	96,4	94,6	88,4	92,0
Průměr	74,7	79,0	82,3	91,6	84,8	90,5

Pro ověření, že naše metoda nezávisí na konkrétních výcvikových datech, byl proveden další test náhodným výběrem vzorků pro tři testy. Každý test byl proveden pro každou podmnožinu 200krát a průměrný výkon (průměrná přesnost \pm směrodatná odchylka) byla vypočtena, viz Tab. 3. Pro první a druhý test byla zjištěna průměrná míra rozpoznání u podskupin srovnatelná s výsledky uvedené v Tab. 2. Naše metoda však dosáhla celkové míry rozpoznání 80 %.

Tab. 3 Míra rozpoznání vertikálních a standardních odchylek (%) naší metody pro datovou sadu MSR-action3D v náhodných testech

	První test	Druhý test	Křížový test
AS1	97,4 \pm 0,9	98,5 \pm 1,1	84,8 \pm 4,4
AS2	96,1 \pm 1,5	97,8 \pm 1,4	67,8 \pm 4,3
AS3	97,7 \pm 1,2	98,9 \pm 1,1	87,1 \pm 3,7
Průměr	97,1 \pm 1,2	98,4 \pm 1,2	79,9 \pm 4,1

V této metodě jsou čtyři hlavní komponenty: generování hloubkové mapy (pro tři pohledy) pro každý rámec, generování vlastností DMM, snížení rozměrů (PCA) a rozpoznání akcí. Náš kód je napsán v Matlabu a testován na počítači s procesorem Intel Core I7 s frekvencí 2,67 GHz a 4 GB RAM. Průměrná doba zpracování každé složky je uvedena v Tab. 4. Průměrný počet snímků v dané sekvenci je přibližně 30. Z těchto časů můžeme usoudit, že je tuto metodu možné použít v reálném čase.

Tab. 4 Průměrná a standardní odchylka času zpracování komponent naší metody

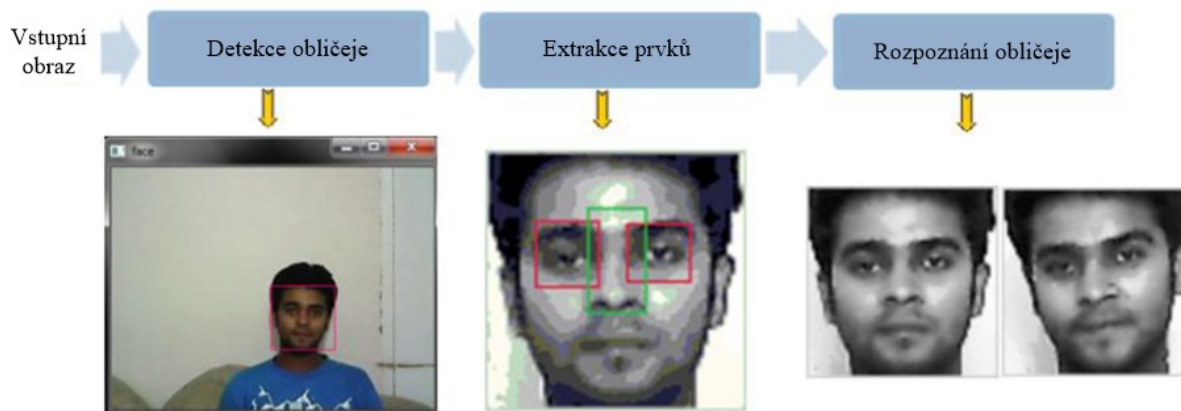
Složky	Doba zpracování [ms]
1	2,0 \pm 0,4 / snímek
2	3,3 \pm 0,6 / snímek
3	2,5 \pm 1,2 / sekvence
4	1,8 \pm 0,5 / sekvence

2.4 Rozpoznání tváře pomocí Lokálního binárního vzoru (LBP)

LBP je texturní algoritmus pro rozpoznávání obličejů, který popisuje strukturu a tvar digitálních obrazů. Předběžně zpracovaný nebo obličejový snímek je nejprve rozdělen do malých bloků, ze kterých jsou vytvářeny histogramy LBP a pak jsou spojeny do jediného vektoru. LBP je definována

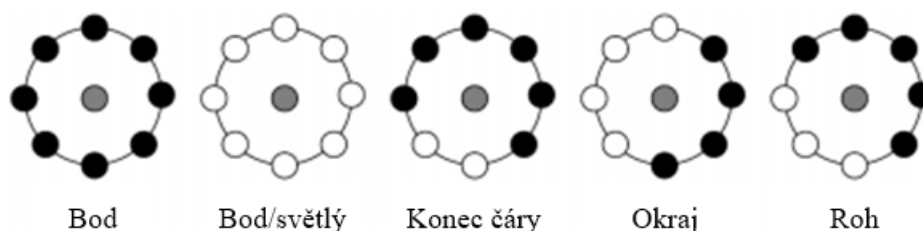
jako uspořádaná sada binárních porovnání intenzity pixelů mezi středovým pixelem a jeho okolními pixely.

K identifikaci tváře pomocí systému rozpoznávání obličeje digitální obraz prochází třemi hlavními fázemi, tj. Detekcí obličeje, extrakcí prvků a rozpoznáním obličeje. Blok detekce obličeje odděluje oblast obličeje od zbytku obrázku na pozadí. Ve fázi extrakce prvků jsou extrahovány nejužitečnější a jedinečné prvky obličeje viz. Obrázek 9. Jakmile jsou tyto příznaky získány tak se obraz porovnává se snímky v různých třídách. Obraz s nejmenšími rozdíly vzdáleností s obrazy ve třídách získává nejvyšší hodnocení a můžeme říci, že se jedná o totožný obraz.

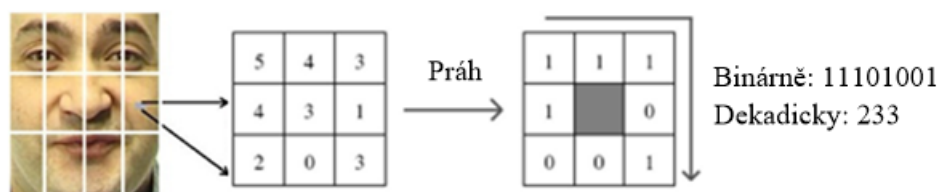


Obrázek 9. Typický systém rozpoznání obličeje

Důležitým zvláštním případem LBP je jednotný LBP. Jednotný deskriptor LBP obsahuje maximálně dva bitové přechody od 0 do 1 nebo naopak. Vzhledem k tomu, že přidělený binární řetězec musí být považován za kruhový, není tedy možný výskyt pouze jednoho přechodu v deskriptoru LBP. To znamená, že LBP nemá buďto žádné přechody nebo má dva přechody. 11111111 a 10001111 jsou příklady jednotných binárních vzorků s nulovými bitovými přechody a dvěma bitovými přechody. Je-li P celkový počet vzorkovacích bodů na okraji kruhu, pak jsou možné kombinace pro vzory se dvěma bitovými přechody vypočítány pomocí $P(P-1)$, což je činí mnohem menší, a proto je snazší s nimi pracovat ve srovnání s nejednotnými vzory, které mají možné kombinace 2^p . Dalším důvodem, proč by mělo být použito jednotné LBP, je to, že detekuje pouze nejdůležitější příznaky v předběžně zpracovaných snímcích, jako jsou rohy, skvrny, hrany a konce čáry, jak je znázorněno na Obrázek 10.



Obrázek 10. Jednoduché texture zachycené LBP operátorem

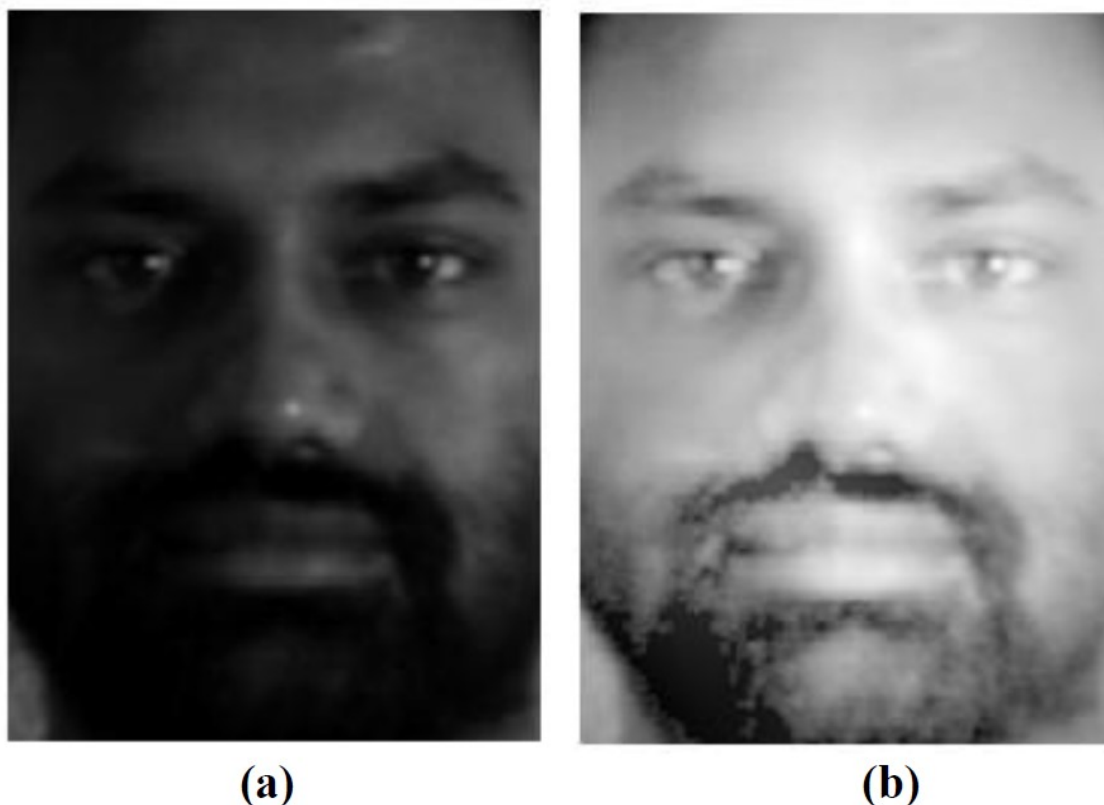


Obrázek 11. Původní LBP operátor

Velikost oblasti – Délka vektoru závisí na bloků, na které byl obraz rozdělen. Například m počet bloků vytvoří vektor m-krát větší.

Matematický modul – Chceme-li použít pole o různých velikostech. Je LBP operátor rozšířen o kružnici o poloměru R od středového pixelu. P je počet vzorkovacích bodů na okraji tohoto kruhu a je porovnán s hodnotou středového pixelu.

Postup je takový, nejdříve se rozdělí obraz na několik malých bloků, z nichž jsou funkce extrahovány (Viz. Obrázek 11). Poté dojde k výpočtu histogramů LBP pro každý blok. Následně se spojí všechny histogramy LBP a získá se jeden zřetěžený vektor. Snímky lze pak porovnat měřením podobnosti (vzdálenosti) mezi histogramy. Problém ovšem nastává s rozdílným osvětlením. Aby se tomuto problému zabránilo je na snímcích použit algoritmus SSR [9], [10]. Digitální snímek se skládá ze dvou komponent, tj. faktor osvětlení (nízkofrekvenční komponenta) a faktor odrazu (vysokofrekvenční komponenta). Abychom odstranili faktor osvětlení používá se Gaussovský filtr, který odstraní nízkofrekvenční komponenta.



Obrázek 12. (a) Snímek před použitím SSR algoritmu, (b) snímek po použití SSR algoritmu

Metoda LBP implementace byla provedena v Matlabu na různých typech snímků pořízených z datového souboru Yale [21]. Při použití algoritmu SSR jsou zaznamenány vyšší míry rozpoznání než

bez jeho použití. Metoda LBP vytvořila vektor pro každý snímek z předzpracovaných snímků a snímků výcvikové sady. Následně byla vypočtena matice obsahující vzdálenosti mezi každou dvojicí snímků. Po normalizaci osvětlení pomocí SSR algoritmu s LBP algoritmu s nastaveným poloměrem 2 pixelu a počtem vzorkovacích bodů na 16, dosahuje tato metoda 100 % míry rozpoznání.

2.5 Rozpoznání snímků obličeje s nízkým rozlišením pomocí CLPM

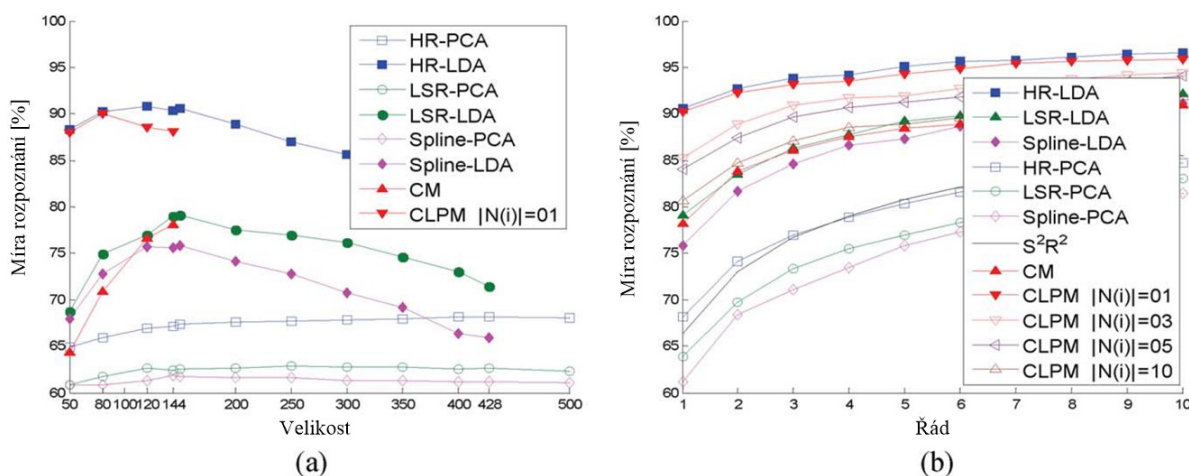
Tato metoda je založená na tzv. Spojených mapách (Coupled mappings – CMs), která promítá snímky s nízkým rozlišením do jednotného prostoru, který je upřednostňován pro klasifikaci. Toto mapování se získá optimalizací cílové funkce, aby se minimalizoval rozdíl mezi shodou (tj. snímky s nižším rozlišením (LR) a jejich protějšek s vysokým rozlišením (HR).

Ve většině případů, úkol rozpoznání obličeje může zjednodušit nalezení vzdálenosti mezi snímky s nízkým rozlišením a vysokým rozlišením. Ovšem některé vzdálenosti, euklidovské vzdálenosti, nemohou být spočteny pouze z velikostí LR a HR snímků z důvodu rozdílné velikosti. Proto by se využila metoda super-rozlišení (SR metoda) k projekci LR snímku do HR prostoru a vzdálenost se vypočte. Nicméně většina existujících SR metod není nejvhodnějším způsobem pro zlepšení kroku rozpoznání. Jelikož SR metoda je náročná výpočty a čas.

Na rozdíl od předchozí metody se tato metoda snaží promítnout body s původními vlastnostmi HR a LR snímků do jednotného prostoru použitím spojených map. Tato metoda rozpoznání obličeje LR snímků má dvě fáze. Jedním z nich je fáze offline, která zahrnuje naučení spojených map a provedení transformací na zachycených HR snímcích. Druhá je on-line fáze, která se skládá z transformace LR snímku a porovnávací funkce. Tato metoda nevyužívá žádnou SR metodu a díky tomu je vhodnější při použití v reálném čase.

Testování metody CLPM bylo provedeno na databázi obličejových snímků FERET [14]. Výcviková část obsahovala 1002 obličejových snímků z frontálního pohledu od 429 osob. Testování proběhlo na standardní galerii s 1196 snímky a na sadě „fařb“ s 1195 snímky. Ve všech experimentech měly HR snímky rozlišení 72x72 pixelů. LR snímky měly rozlišení 12x12 pixelů a byly vygenerovány pomocí operace vyhlazování a podvzorkování.

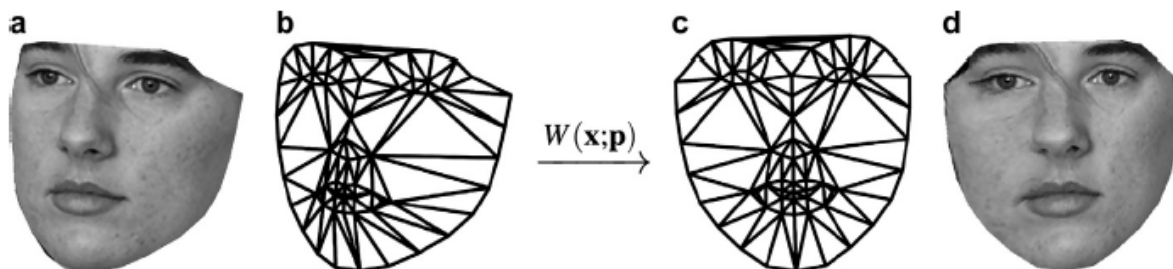
Tato metoda byla porovnána s jinými metodou S^2R^2 (s vlastnostmi PCA) a dalšími dvoukrokovými algoritmy jako HR-PCA/LDA (používající HR snímky), Spline-PCA/LDA a LSR-PSA/LDA. Výsledky a porovnání těchto metod je znázorněno na Obrázek 13.



Obrázek 13. (a) Míra rozpoznání s rozdílnou velikostí snímku, (b) Výsledky hromadného rozpoznání

2.6 Plně automatická normalizace obličeje a jednoduché rozpoznání obličejů bez omezení prostředí

V tomto článku je představena metoda normalizace pózy založené na částečně afinním plánu, který může normalizovat obě změny jak v rovině, tak mimo rovinu. Plán se aplikuje pomocí trojúhelníkových částí, popsanych níže. Celkový proces je znázorněn na Obrázek 14.

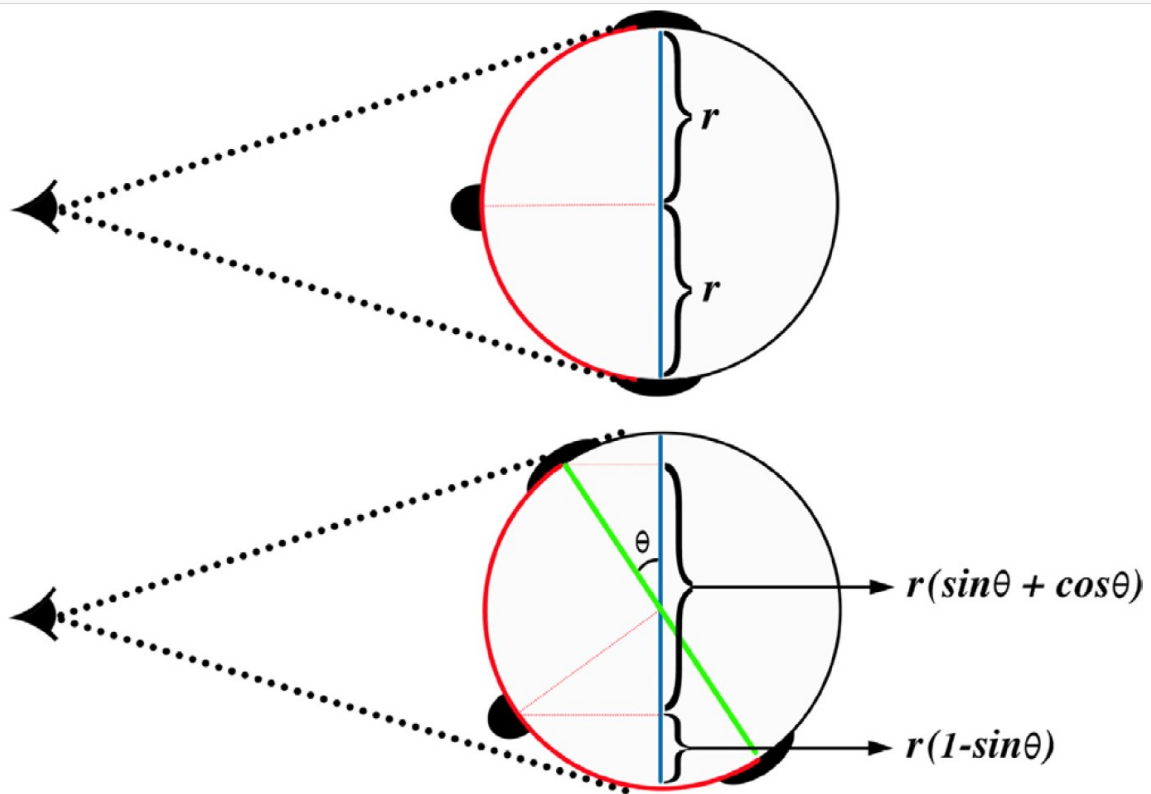


Obrázek 14. Převedení natočeného obličeje do frontální polohy. (a) Natočený obraz obličeje. (b) Nastavení síťové masky odpovídající obrázku otočeného obličeje. (c) Triangulace do základní (čelní) polohy. (d) Převedený snímek obličeje.

Modelování obličeje bylo hlavní součástí aktivního modelu vzhledu (dále AAM – Active appearance model). Vzhledem k modelovým parametrům AAM rekonstruuji konkrétní tvář pomocí statistických modelů tvaru a vzhledu. Parametry modelu se získají maximalizací shody mezi instancí modelu a tváří přizpůsobením AAM snímku vstupního obličeje. Jak je znázorněno na obr. 1, částečný afinní plán $W(x, p)$ přenáší čelní instanci do základního tvaru. Po nastavení AAM má každý trojúhelník v síti AAM odpovídající trojúhelník v čelním pohledu. Při použití souřadnic vrcholů v síti AAM se souřadnice příslušného trojúhelníku v základní síti vypočítají z aktuálních parametrů tvaru.

Výcvik modelu AAM bylo provedeno na 900 pozitivních vzorcích z databáze MULTI-PIE, a 1218 negativních vzorcích z databáze lidí INRIA, která obsahoval snímky z venkovního prostředí bez osob. Tyto vzorky snímky obličeje měli různý úhel natočení od -45° do $\pm 90^\circ$.

Ve fázi testování, protože používáme orientační body pro inicializaci našeho AAM, detekce směsi s 39 vrcholy (orientačními body), odhadujeme umístění zbývajících 29 orientačních bodů na základě informací získaných z topologie obličejových orientačních bodů. Bez ztráty obecnosti, pokud předpokládáme, že pohled shora na lidskou hlavu je kruh s poloměrem r . Obrázek 15 zobrazuje viditelnou oblast levé a pravé strany obličeje z předního pohledu a s natočením.

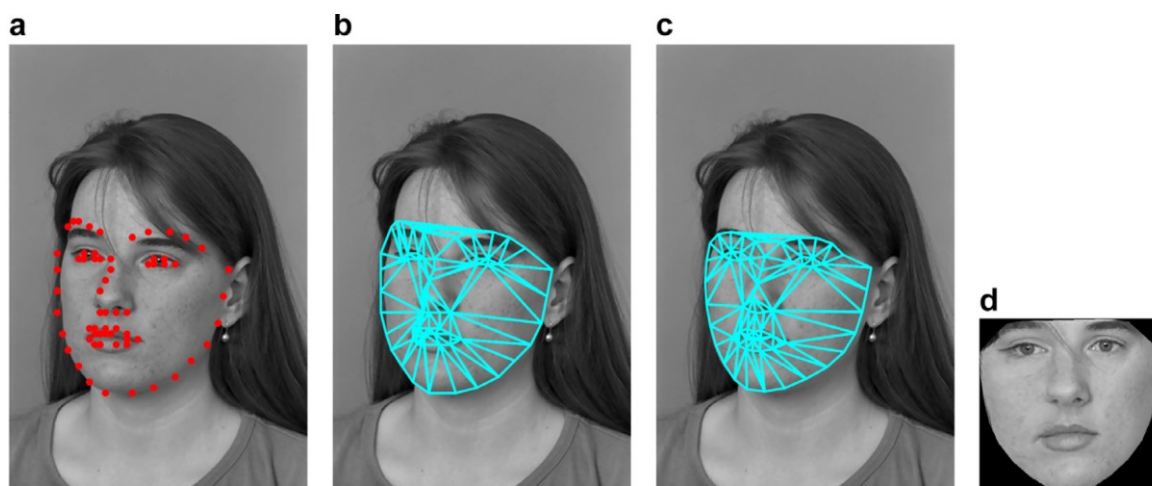


Obrázek 15. Lidská hlava z předního pohledu a s natočením při pohledu ze shora

Jak je znázorněno, poměr mezi viditelnými plochami na obou stranách obličeje je:

$$\gamma = \frac{1 - \sin(\theta)}{\sin(\theta) + \cos(\theta)} \quad (3)$$

V případech, kdy fáze lokalizace orientačních bodů vybírá směs 39 vrcholů, jsou tyto orientační body nastaveny na nejlépe rozeznatelnou polovinu obličeje. Vybraná směs poskytuje odhad uvažovaného úhlu, θ . γ , získané z ekvivalentu, se používá jako měřítko pro hrubý výpočet polohy orientačních bodů na druhé polovině obličeje zrcadlením aktuálních orientačních bodů přes středovou čáru obličeje. Lokalizační algoritmus založený na flexibilní směsi pracuje velmi dobře při hledání obrysu obličeje, ale v detailnějších oblastech, jako jsou oči nebo ústa, není dostatečně přesný. Obrázek 16 (a) znázorňuje výsledek této metody na obrázku vzorového obličeje.



Obrázek 16. Metoda nastavení AMM. (a) Odhadované orientační body. (b) Triangulace počáteční sítě vytvořené časovanými orientačními body. (c) výsledek sestavení. (d) konečný výsledek

V tomto experimentu normalizace snímků obličeje byly použity snímky s rozlišením 120x120 pixelů. Použili se dvě rozdílné metody extrakce příznaků z normalizovaných snímků. A to Gaborovy funkce vln a Histogram orientovaných gradientů (HOG).

Sadu výcviku tvořily tři databáze obsahující 300 snímků obličeje s velkým množstvím různých póz, osvětlení a výrazů. Databáze LFPW obsahovala 1035 snímků získaných prostřednictvím Google, Yahoo a Flickr. Databáze HELEN obsahovala 2330 snímků obličeje získaných z Flickr. Úsměv a neutrální výraz obličeje tvoří většinu výrazů v těchto dvou databázích. Nakonec jsme použili databázi IBUG tvořenou 135 snímky s různými výrazy obličeje. Po sečtení bylo použito 3500 snímků obličeje určené pro výcvik AMM.

Při porovnání s různými metodami na databázi FERET, viz Tab. 5 Porovnání různých metod testovaných na databázi FERET. Můžeme vidět, že tato metoda dosáhla lepších výsledků při malém natočení obličeje $\pm 45^\circ$. Do této hodnoty natočení přináší tato metoda velice dobrou normalizaci, nicméně slabostí této metody byly snímky s větším natočením obličeje. A to konkrétně $\pm 60^\circ$ a více. Při takto vysokých hodnotách natočení docházelo k velké deformaci skryté části obličeje.

Tab. 5 Porovnání různých metod testovaných na databázi FERET

Method	Zarovnání obličeje	Trénování na FERET	bb +60°	bc +45°	bd +25°	be +15°	bf -15°	bg -25°	bh -45°	bi -60°	bj Výraz	bk Osvětlení
LGBP (Zhang et al., 2005)	Automatické	Ne	-	51.0	84.0	96.0	98.0	91.0	62.0	-	-	-
PAN (Gao et al., 2009)	Manuální	Ano	44.0	81.5	93.0	97.0	98.5	91.5	78.5	52.5	-	-
Asthana (Asthana et al., 2009)	Manuální	Ano	32.5	74.0	95.5	98.5	98.0	93.0	87.0	48.0	-	-
Sarfraz (Sarfraz & Hellwich, 2010)	Automatické	Ano	78.0	89.0	97.0	98.6	100	89.7	92.4	84.0	-	-
3DPN (Asthana et al., 2011)	Automatické	Ne	-	91.9	97.0	97.5	98.5	98.0	90.5	-	-	-
CLS (Sharma et al., 2012)	Manuální	Ano	70.0	82.0	90.0	95.0	96.0	94.0	85.0	79.0	-	-
FRAD (Mostafa et al., 2012)	Automatické	Ne	-	82.35	98.47	98.97	100	97.98	87.5	-	-	-
PIMRF (Ho & Chellappa, 2013)	Automatické	Ne	-	91.5	96.5	98.5	98.0	97.3	91.0	-	-	-
PAF (Yi et al., 2013)	Automatické	Ne	93.75	98.0	98.5	99.25	99.25	98.5	98.0	93.75	-	-
FAR (Sagonas et al., 2015)	Automatické	Ne	-	96.0	100	100	100	99.0	96.5	-	-	-
Navržená metoda	Automatické	Ne	91.5	96.0	100	100	100	100	99.0	93.0	99	100

3. Praktická část

3.1 3 Design systému pro detekci a klasifikaci obličejových příznaků

Hlavním cílem bakalářské práce je navrhnout, naprogramovat algoritmus a aplikaci pro detekci a rozpoznání obličeje. Tato část bude obsahovat teoretický návrh algoritmu, jeho realizaci, testování a zhodnocení výsledků pro jednotlivé snímky obličeje.

3.2 Zde by se hodila nová podkapitola: Analýza testovacích dat

Snímky obličeje určené k vývoji a testování algoritmu byly získány z volně dostupné databáze Stirling [9], která obsahuje 312 obličejových snímků obsahující snímky s natočením obličeje o 0° , 45° a 90° s rozdílnými výrazy v obličeji. Databáze obsahuje snímky osmnácti mužů a osmnácti žen po třech a devíti snímcích. Snímky byly pořízeny za stejných světelných podmínek na tmavém pozadí. Obsahovaly snímky 17 mužů a 18 žen, po 3 až 9 snímcích. Databázi bylo nutné, z důvodu rozdílného rozlišení pro jednotlivé snímky normalizovat. Byly odstraněny subjekty, které obsahovaly méně jak 9 snímků, a upraveno rozlišení na 200x255px tak aby byla zajištěna co největší podobnost s původními snímky. Snímky z databáze jsou v úrovních šedi. Tato databáze byla vybrána z důvodu nižších nároku na výpočty. Dále byla databáze rozdělena na tréninkovou a testovací sadu. Testovací sada obsahovala tři snímky každého subjektu, vždy jeden snímek s natočením 0° , natočením 45° a natočením 90° . Tréninková sada pak obsahovala zbývající snímky. Dále byly jednotlivé subjekty tréninkové sady rozděleny do složek. Každá složka obsahovala vždy šest snímků. Příklady snímků v databázi nalezneme na Obrázek 17.

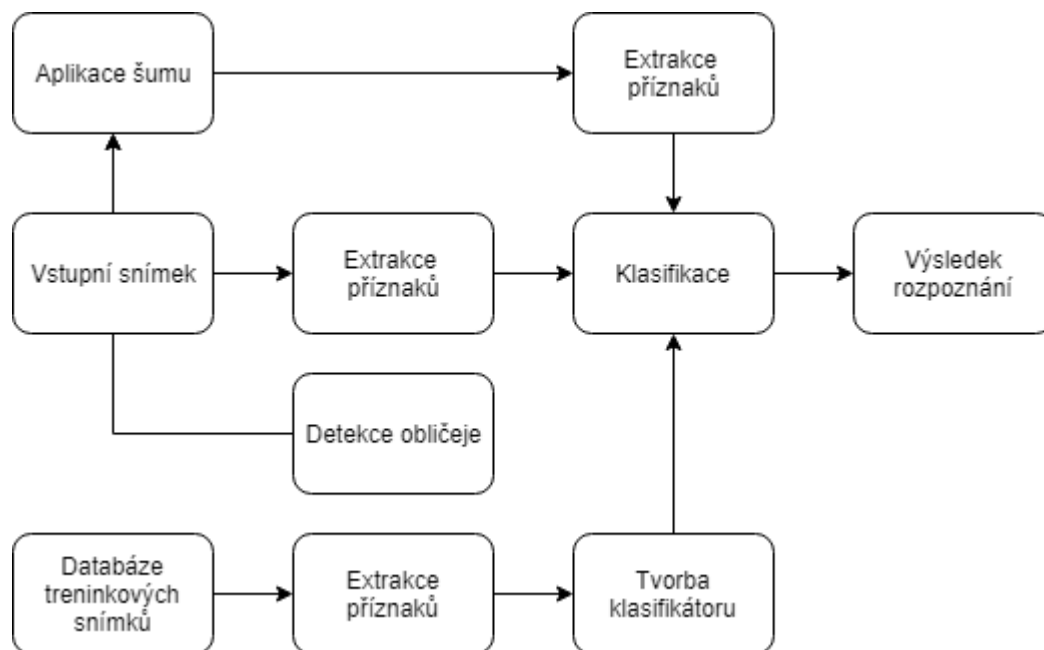


Obrázek 17. Příklady snímků z upravené databáze Stirling

3.3 Návrh algoritmu pro segmentaci a klasifikaci obličejových příznaků

Algoritmus byl navržen tak aby účinně zaznamenal obličejové příznaky ze snímků s nízkým rozlišením a následně je dokázal porovnat s extrahovanými příznaky z tréninkové sady. Skládá ze dvou částí: části tréninku a části testování. Tréninkovou část tvoří určité množství snímků, ze kterých dochází k extrakci příznaků jednotlivých snímků. Na základě těchto extrahovaných příznaků a dalších informací o snímku, jako například umístění ve složce, se vytvoří klasifikátor. Princip testovací části je stejný, pouze s tím rozdílem, že extrakce příznaků probíhá pouze pro jeden snímek. Součástí

testovací části je i aplikace šumu pomocí šumového generátoru a také detekce obličeje na snímku. Konečným postupem je klasifikace. Klasifikace probíhá na základě porovnání extrahovaných příznaků testovaného snímku s klasifikátorem. Výsledkem je čelní snímek osoby z tréninkové databáze, pro kterou byla nalezena největší podobnost.



Obrázek 18. Blokové schéma navrhovaného algoritmu pro analýzu faciálních příznaků

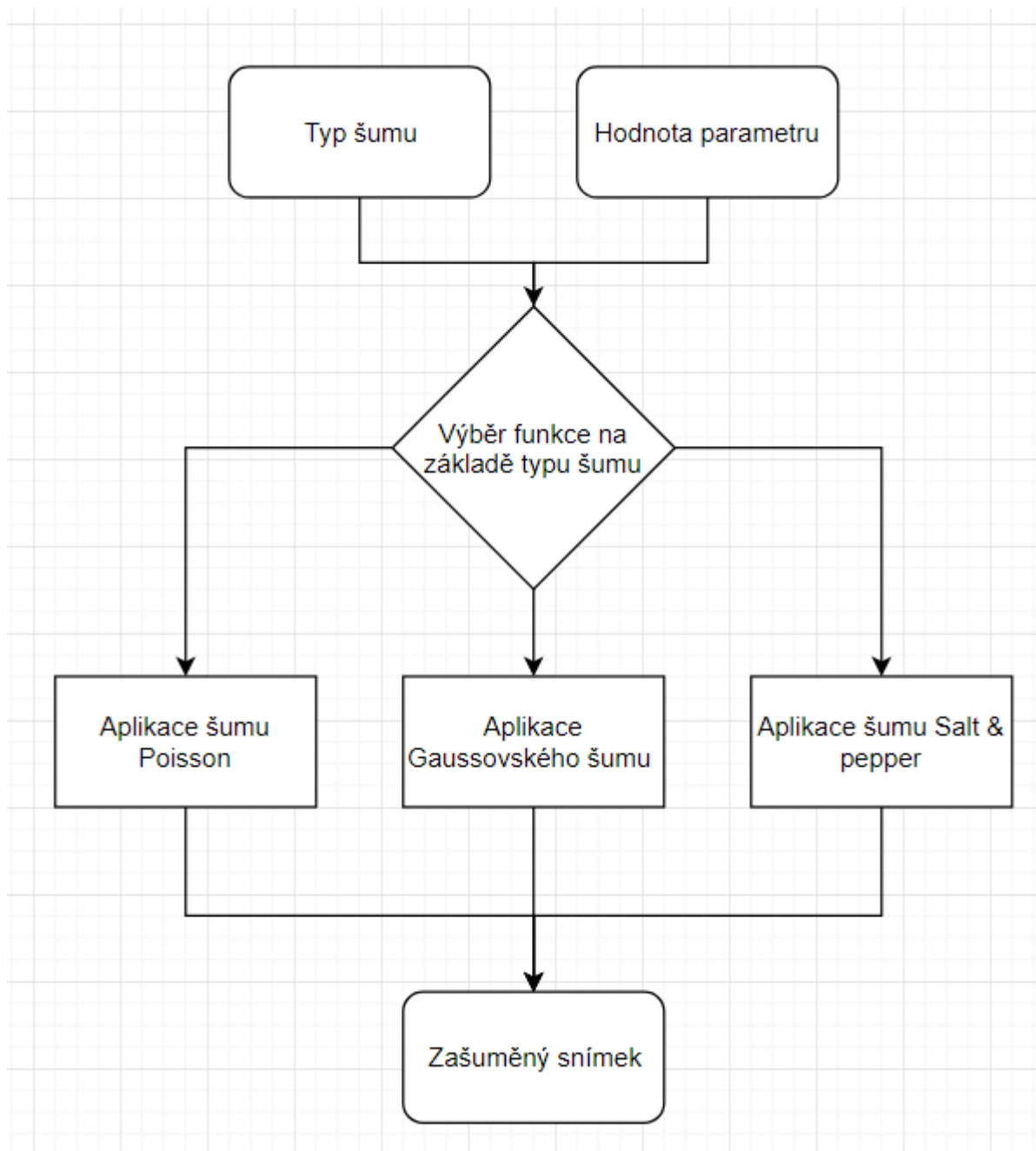
3.3.1 Generátory syntetického šumu

Z hlediska matematických algoritmů, které se využívají pro modelování a klasifikaci faciálních příznaků je kromě samotné efektivity rovněž důležitá robustnost těchto procedur. V tomto ohledu je důležitá analýza, kdy se obecně předpokládá, že faciální charakteristiky jsou významně narušeny šumem. Z hlediska reálných podmínek nelze vždy zajistit unifikované osvětlení a jiné parametry, které vždy předurčují vysokou efektivitu segmentačního a klasifikačního procesu. Bylo použito generátoru šumu, který je součástí programu MATLAB. A bylo použito tří typů šumu: Gaussovský šum, šum Salt & pepper a šum Poisson.

Gaussovský šum je dán svou střední hodnotou μ a rozptylem σ . Jedná se o bílý šum, to znamená, že jeho míra rozložení je normální neboli je stejná na všech pixelech.

Šum Salt & pepper je dán svou denzitou. Denzita určuje kolik procent pixelů snímku bude nahrazeno bílou či černou.

Posledním šumem je šum Poisson. Tento šum je speciálním typem šumu. Distribuce tohoto šumu závisí na datovém typu vstupního snímku. Například pokud se jedná o datový typ double pak jsou hodnoty pixelů vstupního snímku interpretovány jako střední hodnota poissonu zvětšená o e^{-12} . Pokud má například vstupní pixel hodnotu $5,5e^{-12}$ bude výstupní pixel vygenerován z distribuce poissonu se střední hodnotou 5,5 a pak zmenšen o e^{-12} .



Obrázek 19. Diagram algoritmus pro aplikaci šumu

Nejprve je potřeba zvolit typ šumu a jeho parametry. Na základě zvoleného šumu, tedy hodnoty skryté pod textem, se zvolí daná funkce, která aplikuje šum na testovací snímek. Šum Poisson nevyužívá parametru je to z důvodu popsaného v této kapitole výše.

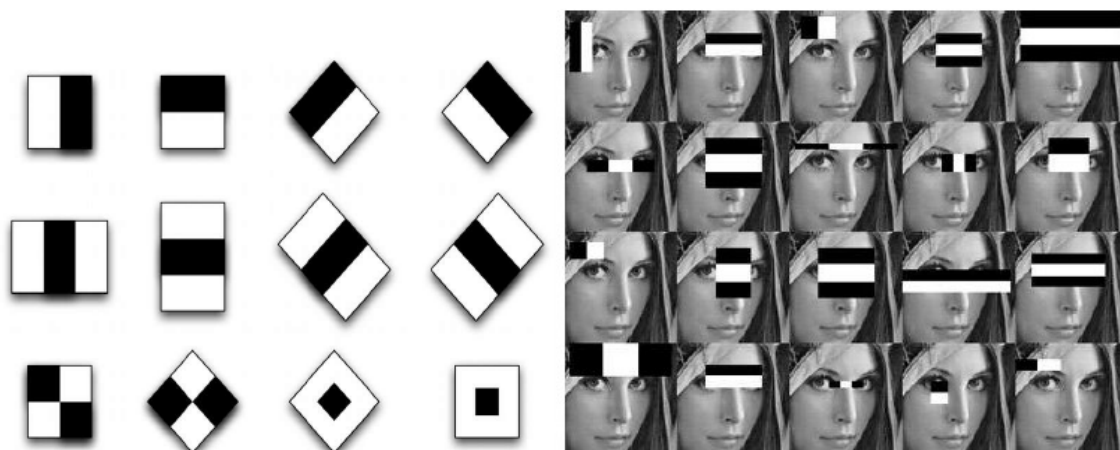
3.3.2 Detekce obličeje na základě Viola-Jones algoritmu

Pro detekci obličeje je použito kaskádového objektu. Algoritmus používá tzv. Haar příznaky viz. Obrázek 20. Tyto příznaky využívají součet jasů jednotlivých pixelů po aplikaci hodnot Haarových příznaků. Jednotlivé bílé pole Harrových příznaku značí hodnotu -1 a černé pole hodnotu +1, tyto hodnoty jsou následně aplikovány na pixely a sumovány. Důvodem tohoto postupu je hledání

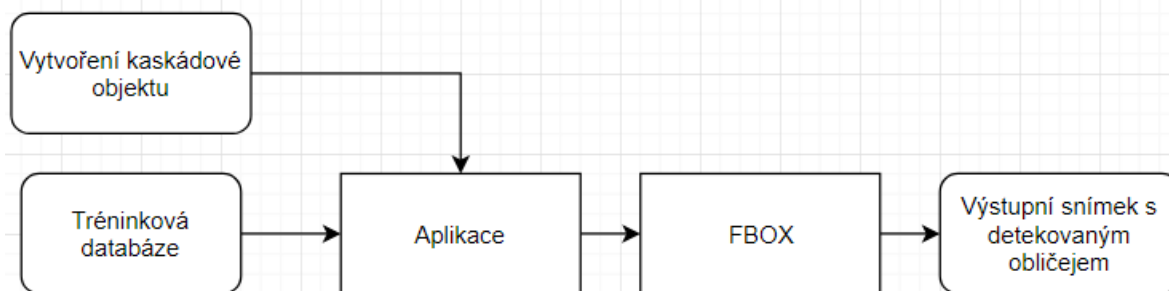
přechodů jasu a podobnosti s Harovými příznaky. Některé příklady podobnosti můžeme vidět na Obrázek 20. Tyto příznaky jsou použity na integrální obraz. Takovýto integrální obraz je daleko menší, než obraz původní a je vypočten podle vztahu:

$$ii(x,y) = \sum i(x,y) \quad (4)$$

Kde $ii(x,y)$ jsou hodnoty pro integrální obraz a $i(x,y)$ pak hodnoty pro obraz původní, hodnoty x a y určují souřadnice pixelů. Přestože integrální obraz je daleko menší získáme velké množství dat po tyto příznaky. Pro redukci těch to dat využívá algoritmus Adaboost (Adaptive boosting). Adaboost využívá kombinaci slabých klasifikátorů, tyto klasifikátor se potom spojují do jednoho silného klasifikátoru. Nicméně tento proces je i tak zdlouhavý, a tedy je tento algoritmus řešen jako kaskádní. Celá tato myšlenka je založena na tom, že většina oken je negativních (neobsahují tedy žádný objekt – typicky pozadí) a pouze menší část je možno klasifikovat jako pozitivní. Teprve při vyhodnocení jako pozitivní použije algoritmus rozšířenou sadu klasifikátorů. V každém kole je tedy snaha vyjádřit co nejvíce negativních podoken.[9] Tento proces je již proveden a detekce obličejů z mých jednotlivých testovacích snímků probíhá jen porovnáváním vytvořeného klasifikátoru.



Obrázek 20. Haarovy příznaky příklady, podobnosti Haarových příznaků s obličejem

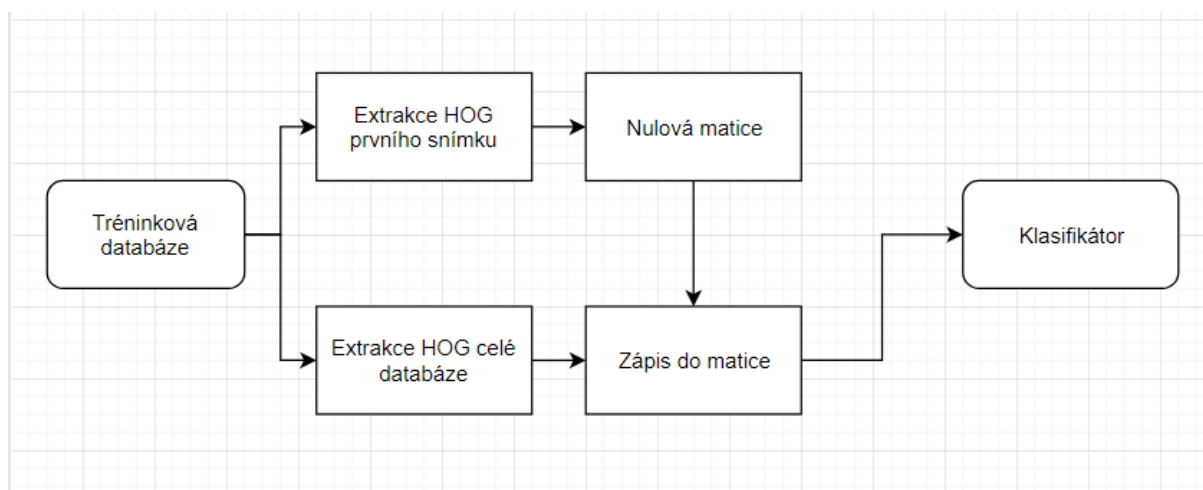


Obrázek 21. Diagram algoritmu pro detekci obličeje

V prvním kroku se vytvoří kaskádový objekt. Tento objekt je aplikován na testovací snímek. Výstupem této aplikace je matice 1×4 kde čtyři hodnoty obsažené v matici symbolizují pixely snímku ohraničující obličej.

3.3.3 Extrakce faciálních příznaků

Jedná se o redukci dimenzionality, která efektivně zastupuje jednotlivé body zájmu obrazu jako kompaktní vektor. Náročnost tohoto kroku se odvíjí podle velikosti databáze a rozlišení snímků. Bylo použito extrakci tzv. příznaků HOG (Histogram of Oriented Gradients – Histogram orientovaných gradientů). Vstupní hodnota této funkce je obraz RGB dán maticí $M \times N \times 3$ nebo obraz v úrovních šedí dán maticí $M \times N$. Histogram orientovaných gradientů je deskriptor příznaků používaný pro detekci a rozpoznání obličeje. Tato metoda počítá s gradienty v určitých částech obrazu.



Obrázek 22. Diagram algoritmus po extrakci příznaků snímku

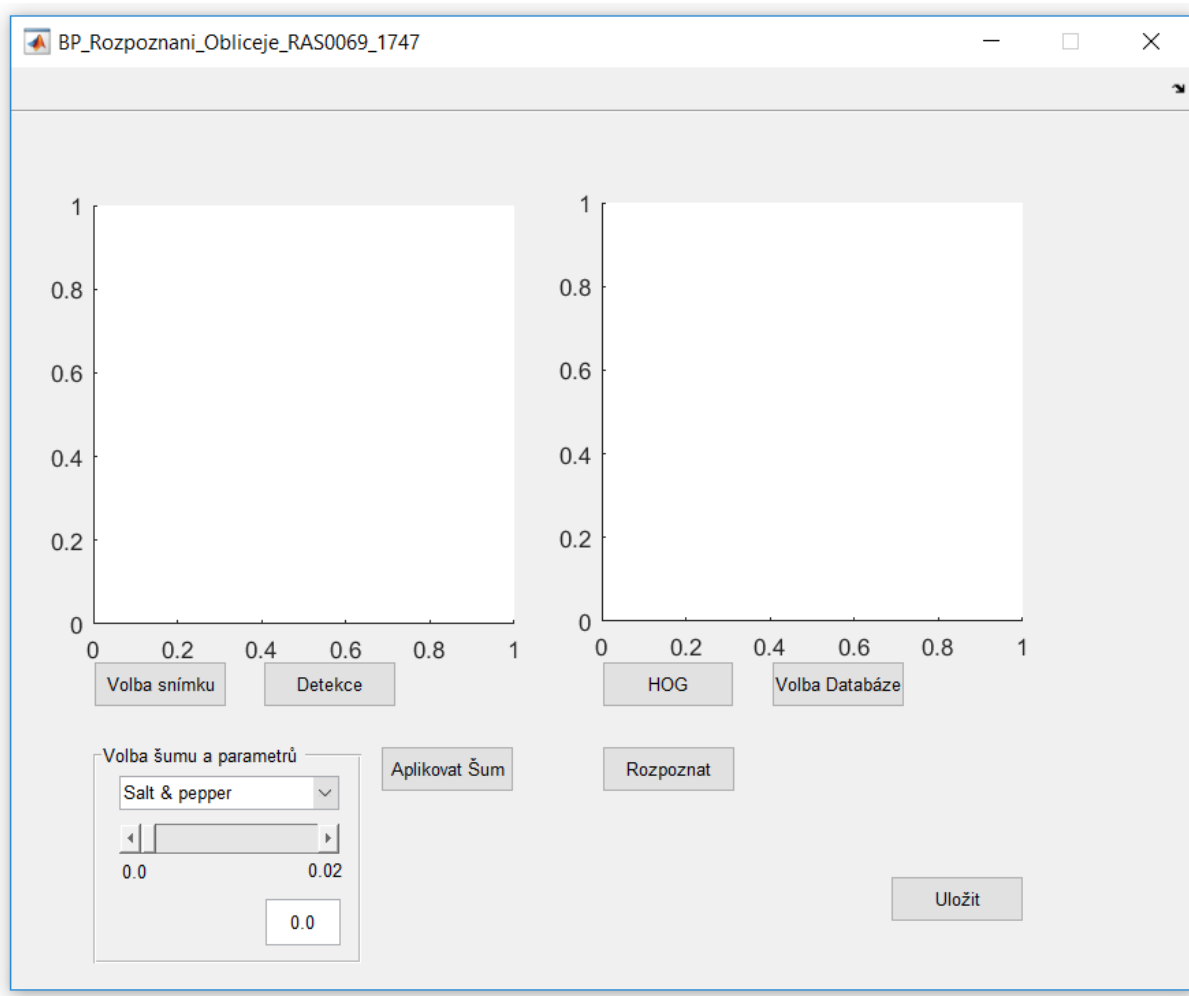
Nejprve dojde k extrakci HOG prvního snímku. Na základě počtu extrahovaných vektorů se vytvoří nulová matice. Do této matice jsou následně zapsány řádek po řádku vektory jednotlivých snímků v trénovací databázi. Ke konci se tato matice použije pro vytvoření klasifikátoru, který slouží pro porovnání s HOG testovacího snímku a nalezení shody.

3.3.4 Tvorba klasifikátoru

Klasifikátor je vytvořen pomocí funkce `fitcecoc()`. V mém konkrétním případě tato funkce vytvoří pomocí dvou matic, matice obsahující názvy složek a matice obsahující extrahované vektory pro všechny snímky tréninkové databáze. Tento klasifikátor následně slouží pro porovnání matice vektorů testovacího snímku a nalezení shody.

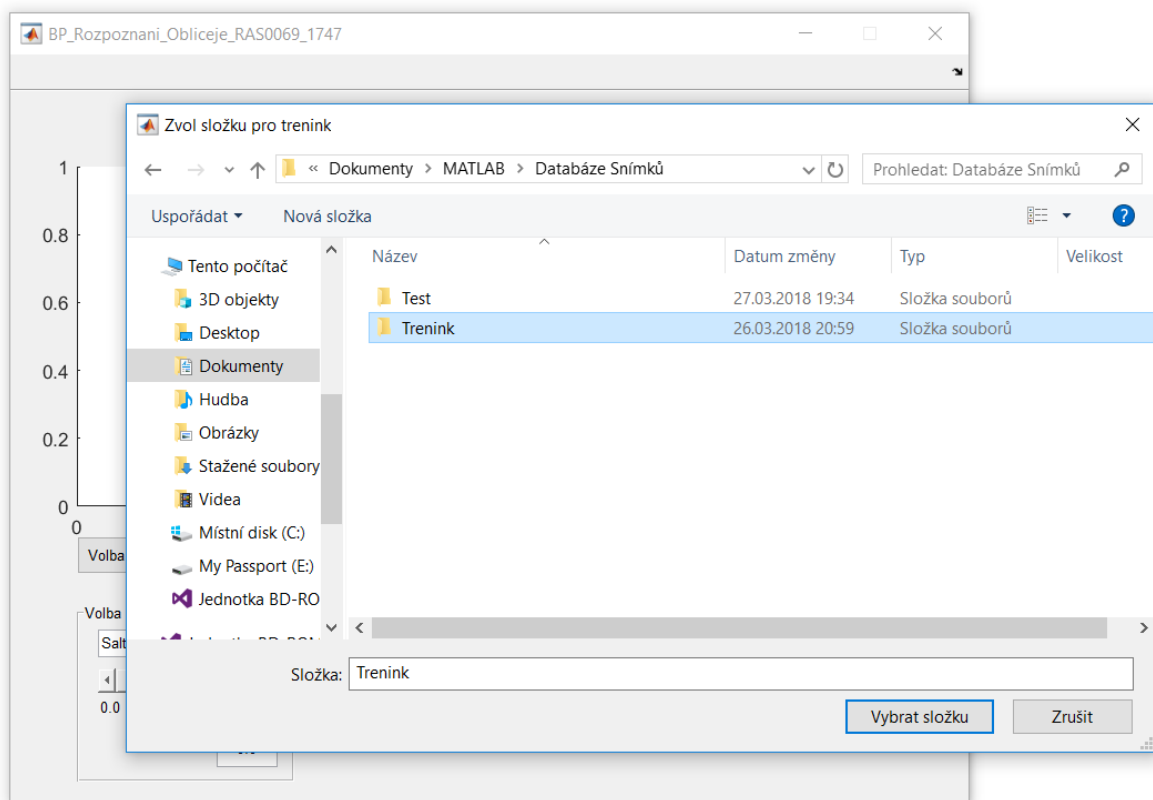
3.3.5 Návrh Graficko-uživatelského rozhraní

Součástí bakalářské práce je i vytvoření graficko-uživatelského rozhraní pro snadnou obsluhu. K vytvoření rozhraní sloužil toolbox GUIDE obsažen v MATLABu. V následující kapitole jsou ilustrační obrázky mého grafického rozhraní s popisky pro jednotlivé prvky rozhraní.



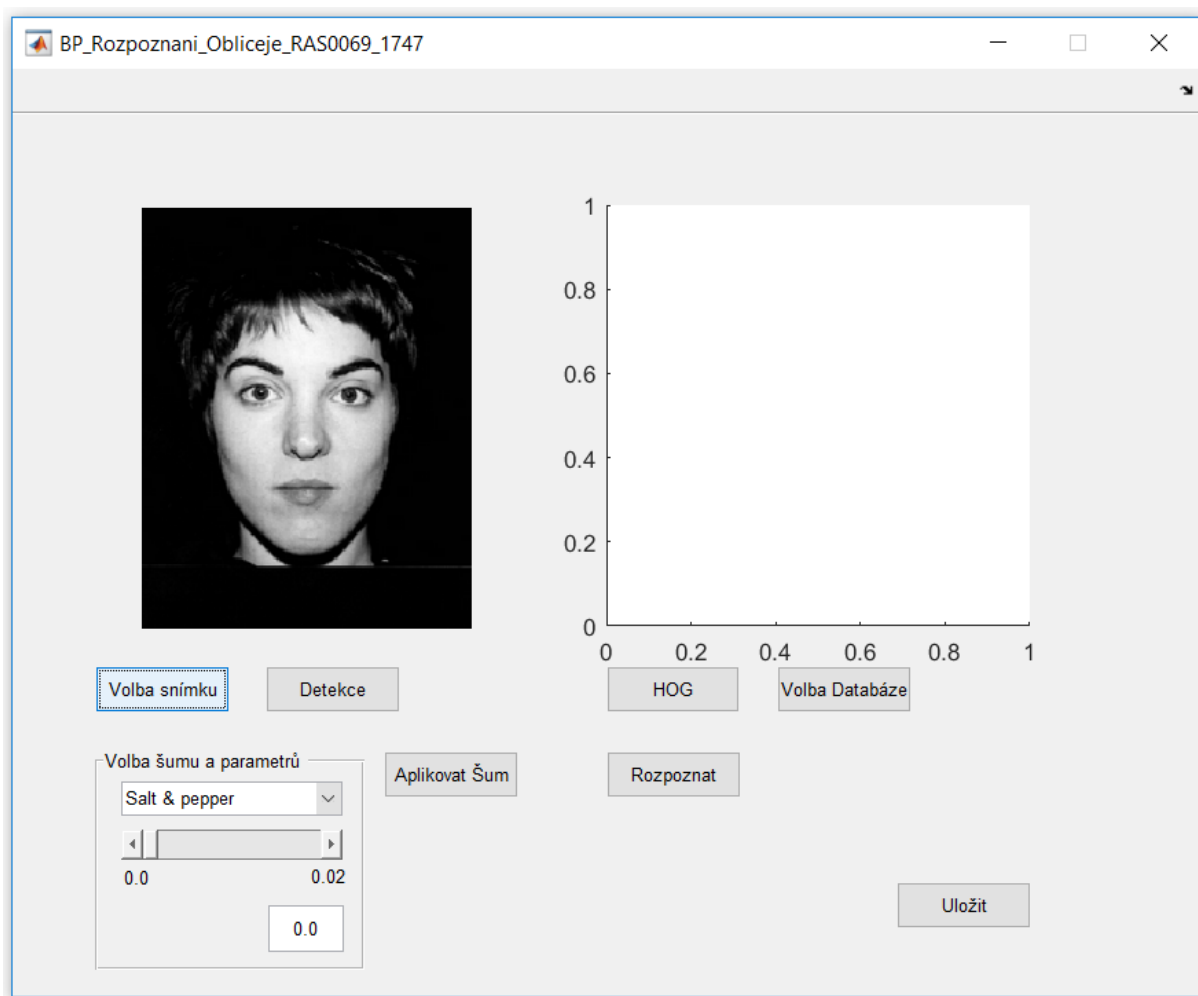
Obrázek 23. Vzhled graficko-uživatelského rozhraní

Nejprve je potřeba načíst složku obsahující tréninkové data. Toho lze docílit pomocí ovládacího prvku „Volba Databáze“. Stisknutím tohoto ovládacího prvku se otevře dialogové okno pro volbu složky obsahující snímky tréninkové databáze. Součástí tohoto kroku je i extrakce matice HOG vektorů pro jednotlivé snímky z databáze a vytvoření klasifikátoru.



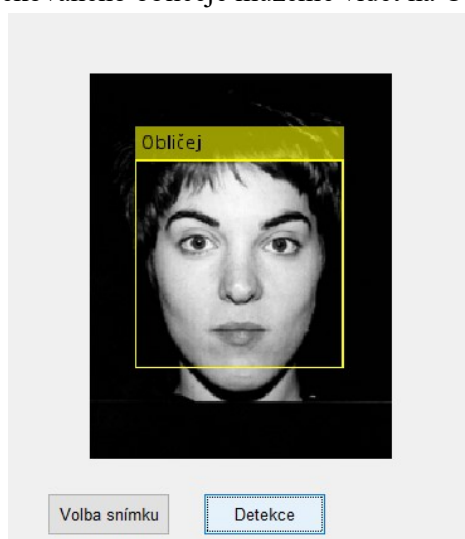
Obrázek 24. Načtení tréninkové databáze

Dalším krokem je načtení testovacího snímku. Tuto akci lze provést pomocí ovládacího prvku „Volba snímku“. Dojde k otevření dialogového okna pro volbu snímku. Uživatel si může zvolit kterýkoliv obrázek z Testovací sady. V případě, že by chtěl snímek jiný, může jej změnit opět pomocí ovládacího prvku „Volba snímku“. Tento snímek se následně zobrazí v levé části okna, jak je vidět na Obrázek 24.



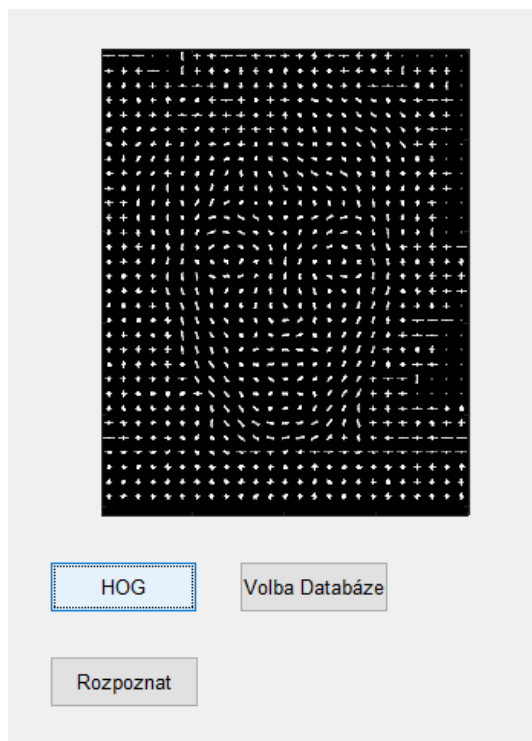
Obrázek 25. Zobrazení testovacího snímku

Pomocí ovládacího prvku „Detekce“ se spustí funkce `vision.CascadeObjectDetector` a provede se procedura popsaná v kapitole 3.3.2. Rámec ohraničující obličej se následně zobrazí na testovacím snímku. Příklad detekovaného obličeje můžeme vidět na Obrázek 26. Detekovaný obličej



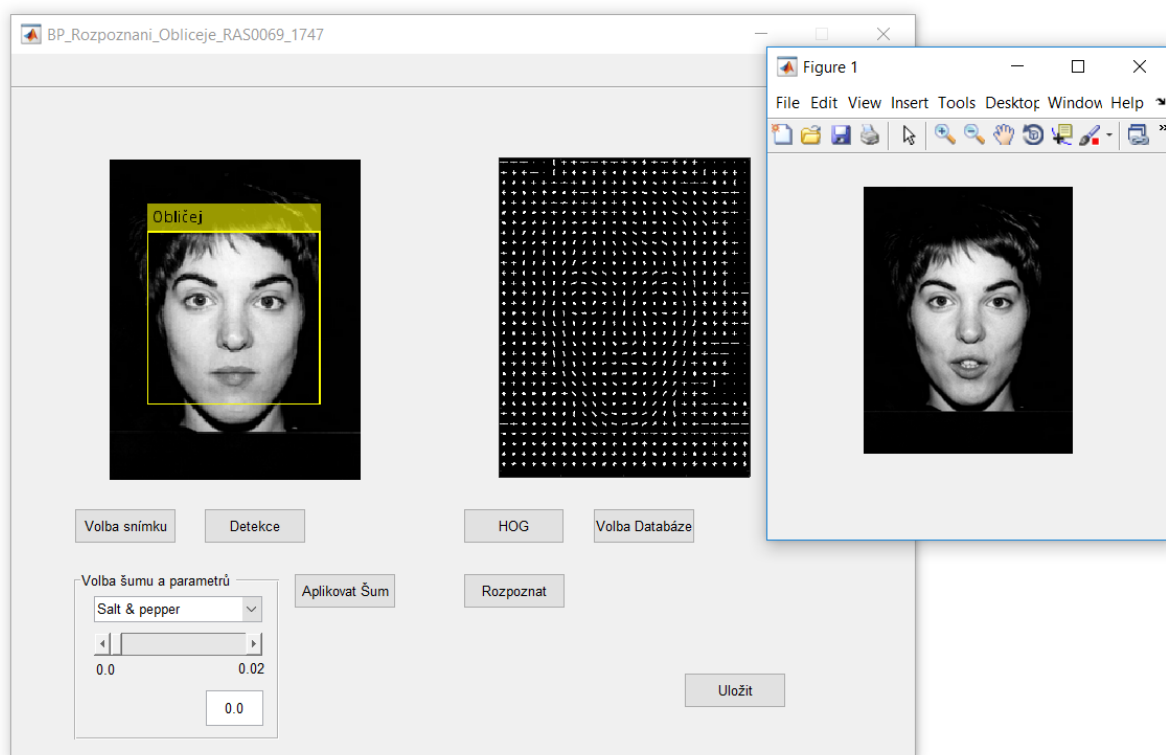
Obrázek 26. Detekovaný obličej

Uživatel následně může pomocí ovládacího prvku „HOG“ zobrazit vizualizaci matice vektorů pro testovací snímek. Tuto funkci může použít i na zobrazení vizualizace pro zašuměný snímek. Vizualizace není závislá na detekovaném obličejí na snímku. Vizualizace se zobrazí vždy pro původní nebo zašuměný snímek. Při porovnání s nativním testovacím snímkem můžeme vidět podobnost.



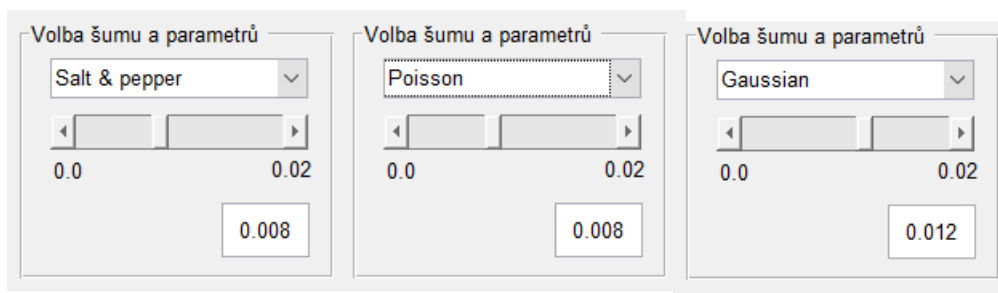
Obrázek 27. Vizualizace extrahovaných HOG vektorů

Funkce skrytá pod ovládacím prvkem „Rozpoznat“ automaticky porovná HOG testovaného snímku s hodnotami HOG v klasifikátoru(modelu). Na základě porovnání zjistí program složku, ve které se nachází nejpodobnější snímek a zobrazí první čelní snímek osoby z tréninkové sady. Tento snímek se zobrazí v novém okně, jak lze vidět na Obrázek 28 vpravo. Tato funkce není závislá na rámci detekovaného obličejí, analyzuje se pouze původní, popřípadě zašuměný snímek.

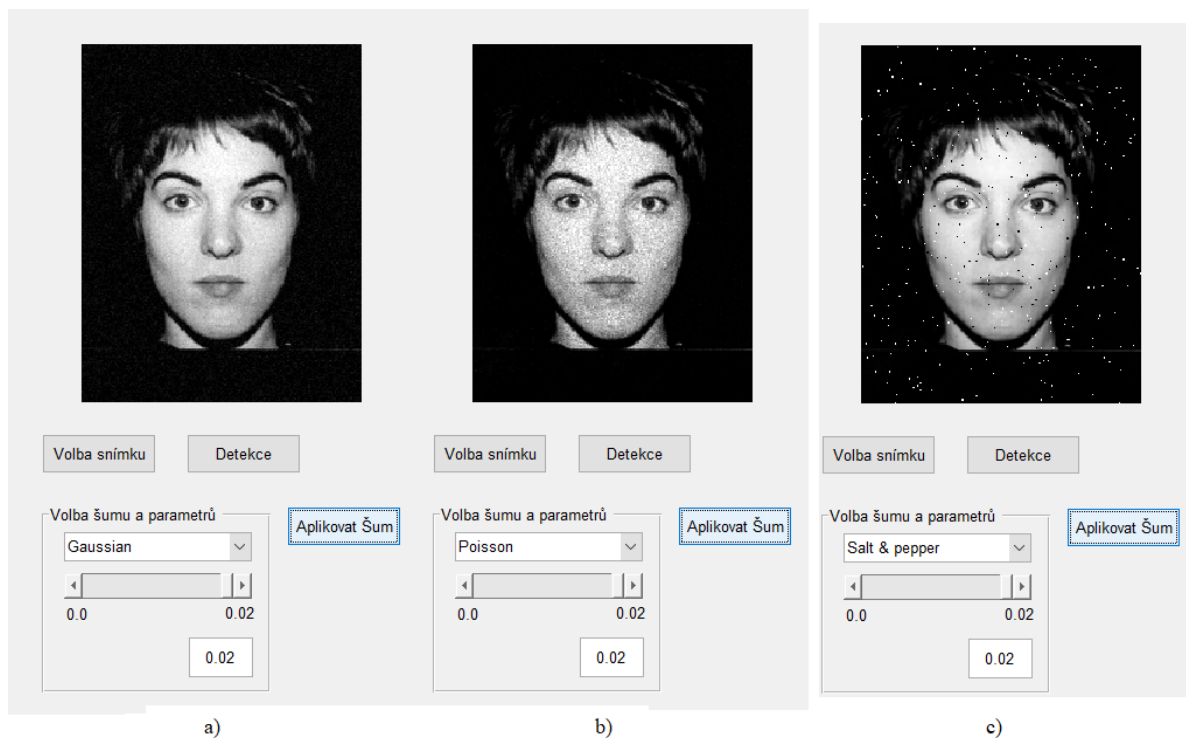


Obrázek 28. Výsledek rozpoznání obličeje

Pro další testování algoritmu je v aplikaci implementování generátor šumu. V okně „Volba parametrů šumu“ může uživatel, pomocí rolovacího menu nahoře, zvolit typ šumu na výběr má tři typy šumu, Gaussovský šum, šum Salt & pepper a šum poisson. Následně buďto pomocí posuvníku nebo editovacího okna je možné nastavit parametry šumu viz. Obrázek 29. Pro šum poisson, který je bez parametru, tento posuvník nemá vliv. Aplikovat šum je možné pomocí ovládacího prvku „Aplikovat šum“. Na základě zvolených parametru dojde k aplikaci šumu na testovací snímek a zobrazení v levé části okna místo původního snímku. Na Obrázek 30. můžeme vidět některé příklady zašuměných dat.



Obrázek 29. Volba šumů a parametrů



Obrázek 30. Zašuměné snímky: a) Gaussovský šum, b) Šum Poisson, c) Šum Salt & pepper

Posledním ovládacím prvkem je prvek „Uložit“. Při stisku tohoto ovládacího prvku dojde k uložení vizualizace HOG vektorů pomocí funkce `imwrite()`.

3.4 Testování algoritmu

Tato část bakalářské práce je věnována testování algoritmu. Algoritmus byl testován pro 32 subjektů s celkovým počtem 96 snímků. Pro každý subjekt testovací databáze obsahovala 3 snímky každý z jednoho pohledu. Ze snímků obsažených v tréninkové databázi byly extrahovány HOG a vytvořen klasifikátor.

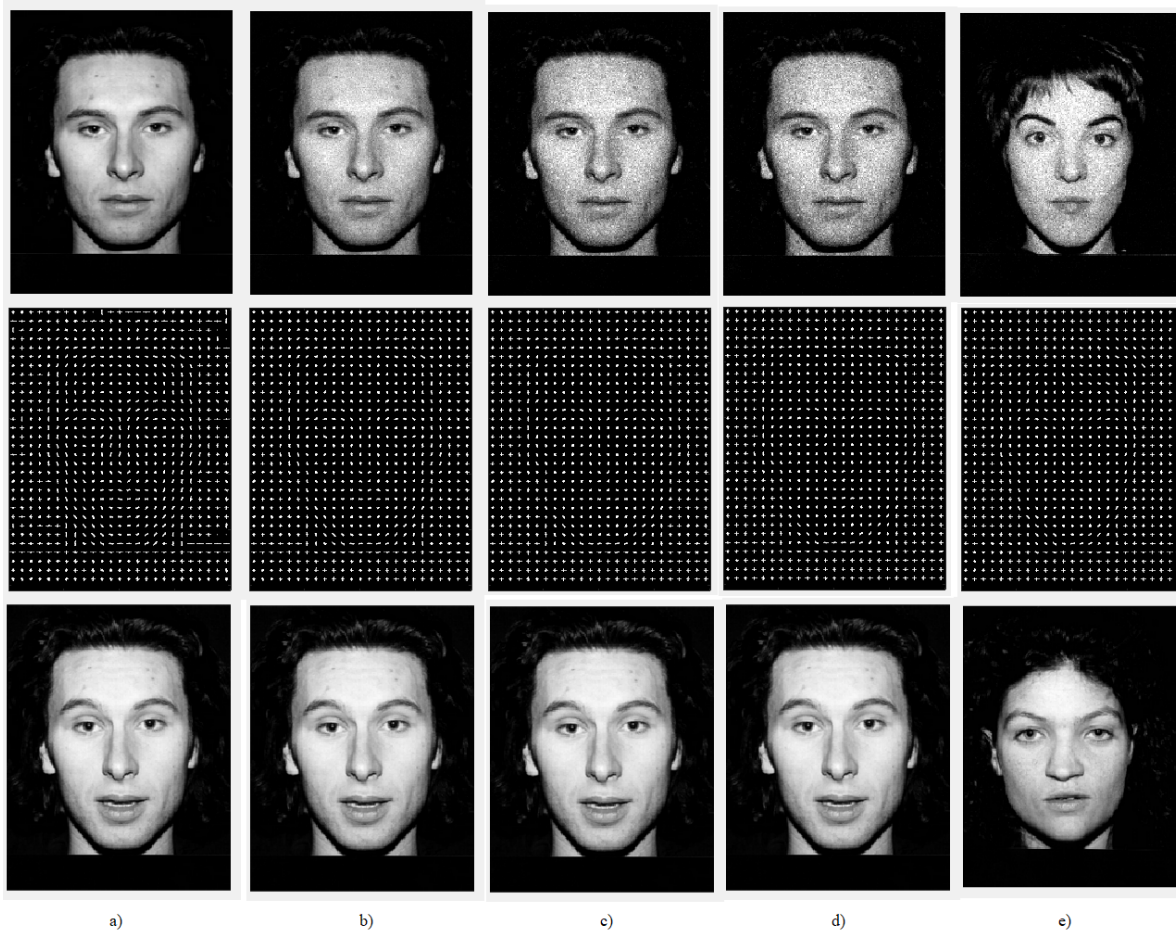
Nejprve proběhne testování algoritmu na nativních snímcích. Na Obrázek 31. jsou některé z výsledků pro různé subjekty s rozdílným natočením. Kde v horní části obrázku jsou testované snímky a v dolní části odpovídající snímky z tréninkové databáze.



Obrázek 31. Příklady výsledků rozpoznání různých subjektů

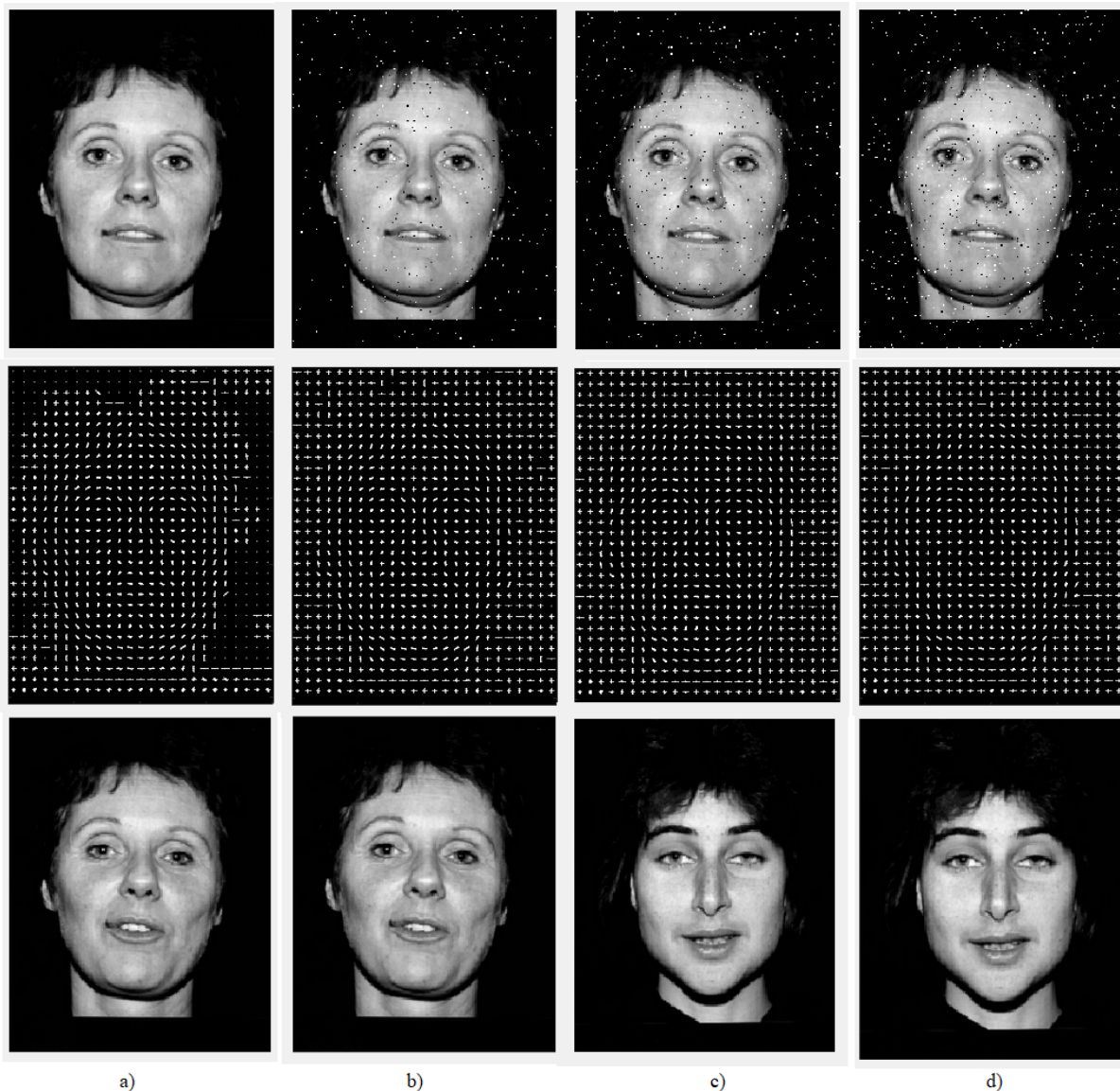
Dalším bodem testování je rozpoznání subjektu při použití šumového generátoru. Byli aplikovány všechny tři typy šumu s rozdílnými parametry. Šum Poisson byl otestován třikrát při rozdílném počtu aplikací, a to jedenkrát, dvakrát a třikrát. Šum Salt & pepper byl aplikován s denzitou $d = 0,01$; $d = 0,015$ a $d = 0,2$. A poslední byl aplikován Gaussovský šum s parametry rozptylu $\sigma = 0,00002$, $\sigma = 0,00005$, $\sigma = 0,0001$ a fixně nastavenou střední hodnotou $\mu = 0,0001$. Tyto parametry byly zvoleny tak aby bylo vidět postupné snížení míry rozpoznání, a tedy aby šel vidět, jak velký má dopad aplikovaný šum.

Testování při rozdílném počtu aplikací šumu Poisson. Kde na Obrázek 32 jsou vidět příklady změn jednotlivých snímků při rozdílném počtu aplikací šumu. Můžeme si povšimnout poměrně razantní změny ve vizualizacích vektorů. Kdy docházelo z počátku převážně ke změnám černého pozadí. Při větším počtu aplikací již docházelo i ke změně rysů obličeje. Na Obrázek 32 a), b), c), d), ale můžeme vidět, že i při velké změně byl subjekt rozpoznán, tyto případy jsou ovšem ojedinělé.



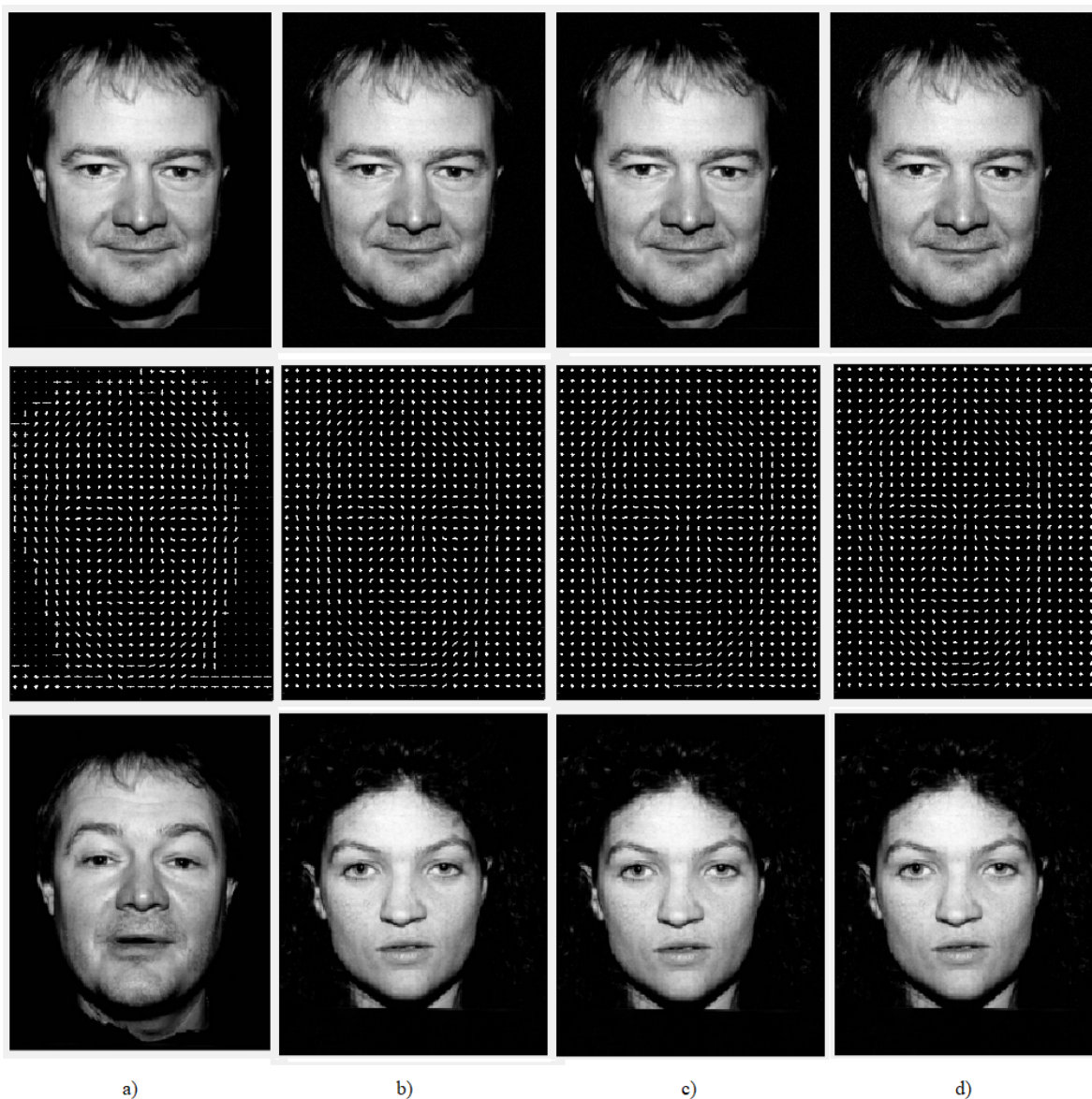
Obrázek 32. Příklady rozpoznání při aplikaci šumu Poisson a) bez šumu; b) 1x aplikován; c) 2x aplikován; d) 3x aplikován; e) 3x aplikován

Testování při aplikaci šumu Salt & pepper. Při aplikaci šumu Salt & pepper dochází k náhodnému nahrazení určitého procenta pixelů snímku za bílé a černé pixel. Přesto je možné tímto simulovat nečistoty na čočce fotoaparátu. Na Obrázek 33. můžeme vidět některé příklady vlivů aplikace šumu.



Obrázek 33. Aplikace šumu salt & pepper a) bez šumu; b) $d = 0,01$; c) $d = 0,15$; d) $d = 0,02$

Posledním testováním byl testování při aplikaci Gaussovského šumu. Při aplikaci šumu byla měněna pouze hodnota rozptylu. Střední hodnota byla fixně nastavena na $\mu = 0,0001$. Gaussovský šum měl na snímky největší vliv. I relativně nízké hodnoty rozptylu výrazně znehodnotily snímek. Hodnoty rozptylu byly počty $\sigma = 0,00002$, $\sigma = 0,00005$, $\sigma = 0,0001$. Na Obrázek 34. Je možné vidět, že i při nejnižších hodnotách rozptylu $\sigma = 0,00002$ snímek nebyl rozpoznán.



Obrázek 34. Aplikace Gaussovského šumu a) bez šumu; b) $\sigma = 0,00002$; c) $\sigma = 0,00005$; d) $\sigma = 0,0001$

3.5 Vyhodnocení výsledků

Tato část bakalářské práce je věnována hodnocení výsledků. Všechny 96 snímků obsažených v testovací databázi bylo testováno, zda pro ně bylo nalezeno správné shody v tréninkové databázi. Testy byly provedeny na nativních datech a také při aplikaci šumového generátoru a třech rozdílných šumů. Pro každý z těchto šumů byly zvoleny tři parametry. Výsledky jsou zaznamenány v grafech pro jednotlivé šумы a nativní data. Tyto grafy zobrazují míru rozpoznání, zvláště mužů a žen, s ohledem na natočení obličeje a aplikovaném šumu.

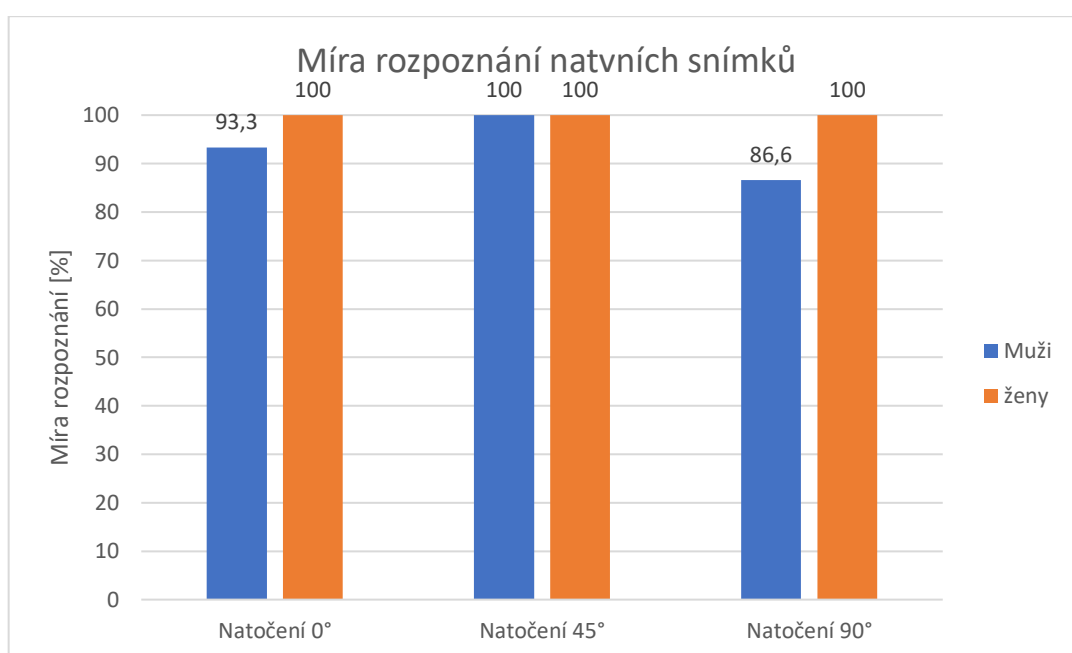
Při testování na nativních datech byla míra rozpoznání vysoká a dosahovala u žen 100 % a u mužů 93,3 %. Nižší procento míry rozpoznání u mužů bylo dáno převážně podobností při natočení 90°. Pouze jeden snímek z čelního pohledu nebyl rozpoznán. Viz. Graf 2.

Při aplikaci šumů byly výsledky o poznání horší. Nejmenší míru rozpoznání vykazoval Gaussovský šum, který i při svých velice nízkých hodnotách rozptylu $\sigma = 0,00002$; $\sigma = 0,00005$;

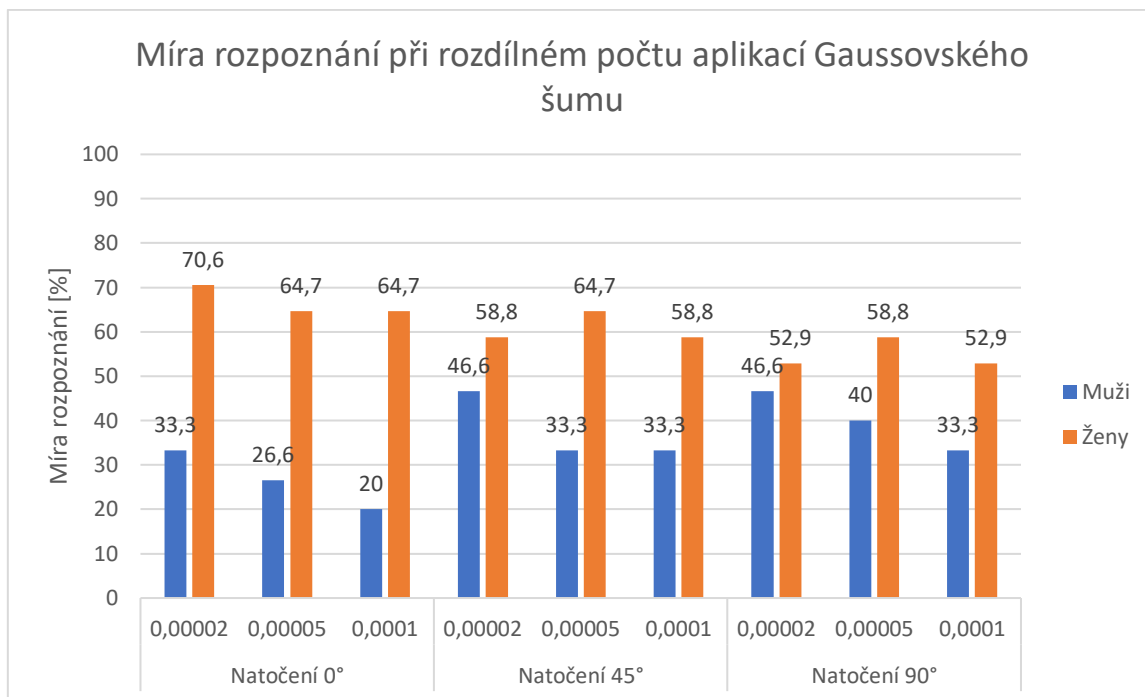
$\sigma = 0,0001$ dosáhl maximální míry rozpoznání pro ženy 70,6 % a nevyšší pro muže 33,3 % pro nejnižší hodnotu rozptylu. Musíme ovšem vzít v potaz to, že hodnoty pro střední hodnotu i rozptyl, které byly použity, jsou velice malé. Při náhodném testování, kdy byly zvoleny hodnoty rozptylu a střední hodnoty desetkrát vyšší již byla míra rozpoznání nulová. Viz. Graf 3.

Testování po aplikaci šumu Salt & pepper s denzitou $d = 0,01$, tato hodnota odpovídá 1 % nahrazených pixelů, se míra rozpoznání pro ženy pohybovala okolo 80 % a pro muže přibližně 55 %. S rostoucí denzitou následně klesala i míra rozpoznání. Viz. Graf 4.

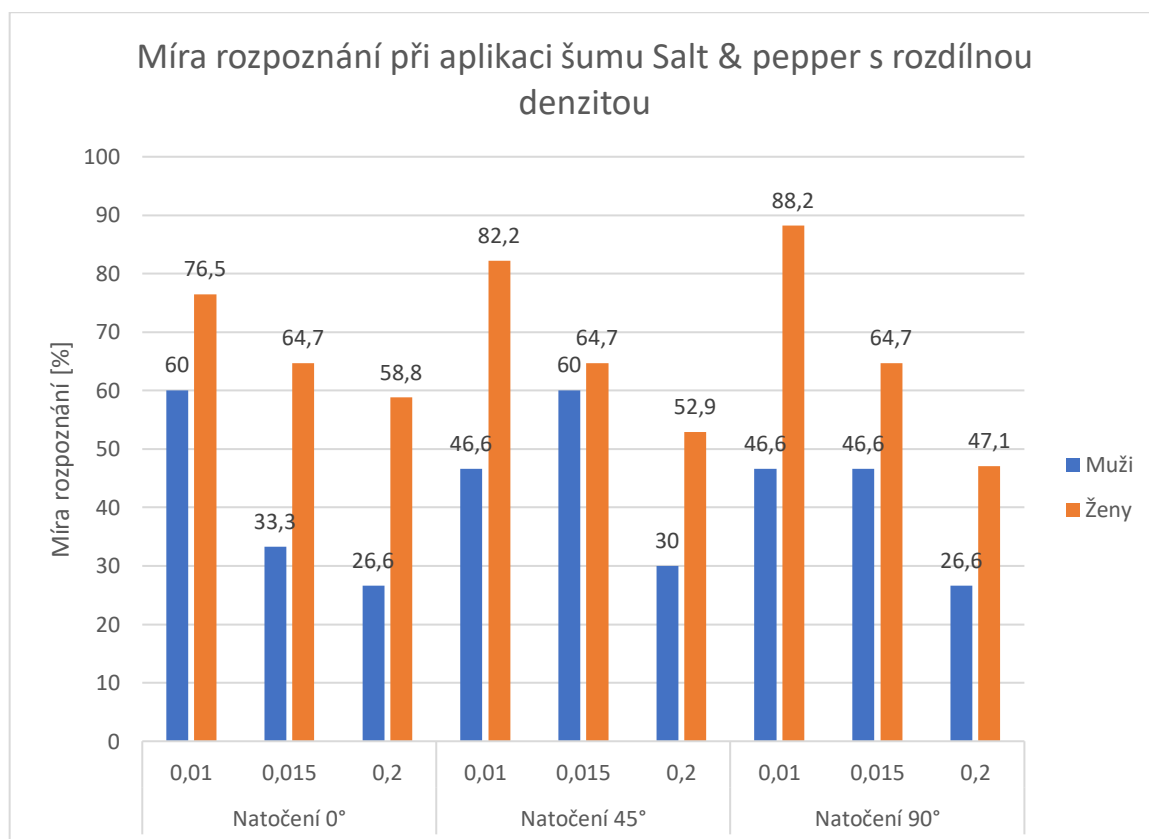
Posledním bodem testování bylo testování při aplikaci šumu Poisson. Míry rozpoznání při jedné a dvou aplikacích byly přibližně podobné ve většině případů se lišily v řádech pár procent. Největší vliv měla vícečetná aplikace šumu na snímky s natočením 90° . U těch to snímků docházelo k poklesu míry rozpoznání u žen až o 30 %. U mužů byl pokles míry rozpoznání přibližně 5 %. Viz. Graf 5.



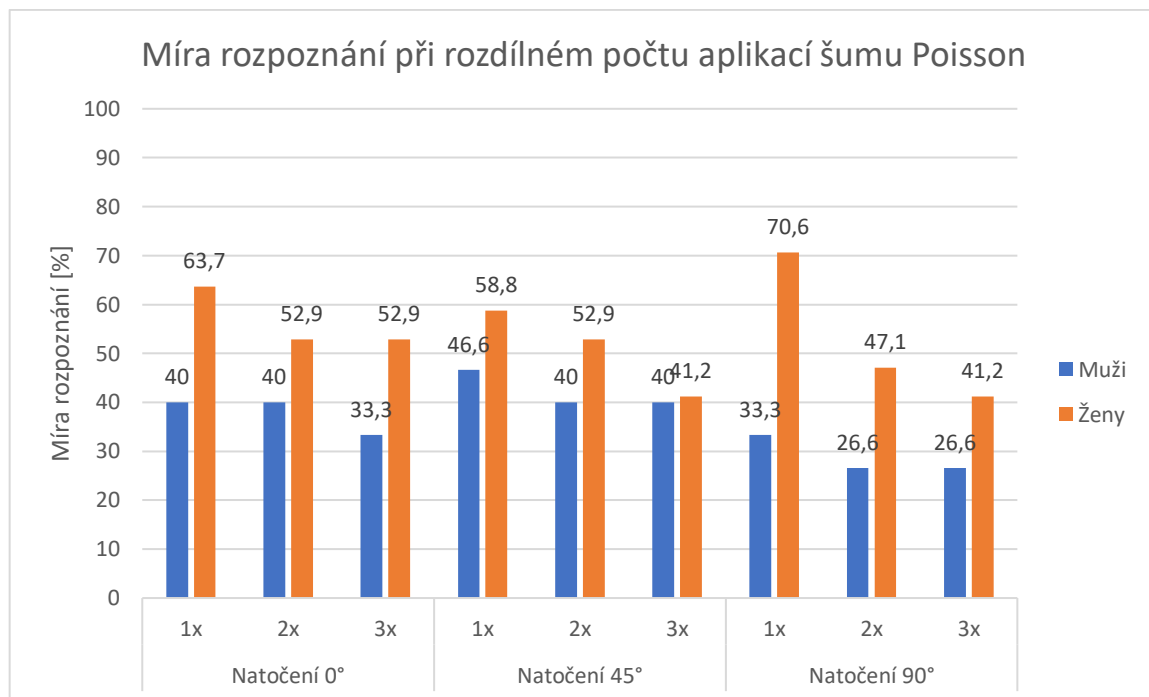
Graf 2. Výsledky testování na nativních datech



Graf 3. Výsledky testování při aplikaci Gaussovského šumu



Graf 4 Výsledky testování při aplikaci šumu Salt & pepper



Graf 5. Výsledky testování při aplikaci šumu Poisson

4. Závěr

Úkolem bakalářské práce bylo navrhnout algoritmus pro detekci obličeje ze snímku a pro rozpoznání uživatele ze snímku. V první části byla vypracovaná rešerše zabývající se metodami pro detekci obličeje a extrakci příznaků ze statického a dynamického obrazu. Díky těmto informacím byl navrhnout a vytvořen vhodný algoritmus, který byl schopen rozpoznat subjekt a nalézt pro něj odpovídající shodu. Finální algoritmus využívá metodu extrakce HOG (Histogram orientovaných gradientů) příznaků pro extrakci obličejových rysů z tréninkové sady a vytvoření klasifikátoru. Tato metoda je dále použita na testovací snímek a dochází k porovnání příznaků s klasifikátorem. Celková databáze obsahovala 288 snímků 15 mužů a 17 žen s rozdílným natočením obličeje a rozdílnými výrazy v obličeji. Tato databáze byla následně rozdělena na dvě sady. Tréninkovou sadu a sadu testování. Testovací sada obsahovala celkově 96 snímků pro všech 32 subjektů, tedy 3 snímky každého subjektu s rozdílným natočením obličeje. Pro usnadnění práce bylo navrženo graficko-uživatelské rozhraní pro usnadnění volby parametrů šumů, detekci obličeje a rozpoznání uživatele ze snímku. Algoritmus vykazoval vysokou míru rozpoznání pro nativní snímky, pro ženy 100 % a pro muže 93,3 %. Pro komplexnější testování algoritmu byly snímky zašuměny pomocí generátoru šumu. Algoritmus byl testován při účinku tří šumu s rozdílnými parametry. Při aplikaci šumu se průměrné výsledky míry rozpoznání pohybovaly okolo 30-63 % jediné pro šum Salt & pepper s denzitou $d = 0.01$ pohybovala míra rozpoznání okolo 70 %. Možným řešením zvýšení míry rozpoznání je extrahovat HOG příznaky pouze výsledného detekovaného obličeje. Tímto by se eliminovalo pozadí a vlasy uživatele a mohlo by být docíleno lepší klasifikace. Tato metoda by pravděpodobně snížila i výpočetní nároky. Výsledná algoritmus je možno nalézt v CD (Příloha 1) přiloženém k bakalářské práci.

Seznam literatury:

- [1] ŠČUREK, Jaromír. *Biometrické metody identifikace osob v bezpečnostní praxi* [online]. Fakulta bezpečnostního inženýrství, Katedra bezpečnostního managementu, Oddělení bezpečnosti osob a majetku, 2008 [cit. 2018-01-05]. Dostupné z: http://www.rucnepsanypodpis.cz/PDF/biometricke_metody.pdf. Studijní text. Vysoká škola báňská – Technická univerzita Ostrava.
- [2] ALI, Ammad, Shah HUSSAIN, Farah HAROON, Sajid HUSSAIN a M. Farah KHAN. Face recognition with Local Binary Pathterns. *Information & Communication Technology* [online]. 2012, (5), 5 [cit. 2018-01-05]. Dostupné z: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.428.5527&rep=rep1&type=pdf>
- [3] LI, Jiadi, Zhenxue CHEN a Chengyun LIU. Low-Resolution Face Recognition of Multi-Scale Blocking CS-LBP and Weighted PCA. *International Journal of Pattern Recognition and Artificial Intelligence* [online]. 2016, **30**(08), 1656005- [cit. 2017-05-28]. DOI: 10.1142/S021800141656005X. ISSN 0218-0014. Dostupné z: <http://www.worldscientific.com/doi/abs/10.1142/S021800141656005X>
- [4] BO LI, HONG CHANG, SHIGUANG SHAN a XILIN CHEN. Low-Resolution Face Recognition via Coupled Locality Preserving Mappings. *IEEE Signal Processing Letters* [online]. 2010, **17**(1), 20-23 [cit. 2017-05-28]. DOI: 10.1109/LSP.2009.2031705. ISSN 1070-9908. Dostupné z: <http://ieeexplore.ieee.org/document/5232898/>
- [5] CHEN, Chen, Kui LIU a Nasser KEHTARNAVAZ. Real-time human action recognition based on depth motion maps. *Journal of Real-Time Image Processing* [online]. 2016, **12**(1), 155-163 [cit. 2017-05-28]. DOI: 10.1007/s11554-013-0370-1. ISSN 1861-8200. Dostupné z: <http://link.springer.com/10.1007/s11554-013-0370-1>
- [6] HAGHIGHAT, Mohammad, Mohamed ABDEL-MOTTALEB a Wadee ALHALABI. Fully automatic face normalization and single sample face recognition in unconstrained environments. *Expert Systems with Applications* [online]. 2016, **47**, 23-34 [cit. 2017-05-28]. DOI: 10.1016/j.eswa.2015.10.047. ISSN 09574174. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0957417415007514>
- [7] MININ, Peter a Yury SHUMILOV. Sharpness estimation in facial images by spectrum approximation. *Signal, Image and Video Processing* [online]. 2017, **11**(1), 163-170 [cit. 2017-05-28]. DOI: 10.1007/s11760-016-0915-4. ISSN 1863-1703. Dostupné z: <http://link.springer.com/10.1007/s11760-016-0915-4>
- [8] PŘINOSIL, Jiří. *Analýza emocionálního stavu na základě obrazových předloh* [online]. Brno, 2008 [cit. 2018-01-05]. Dostupné z: https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=10245. Doktorská práce. Vysoké učení technické v Brně.

- [9] HÖLL, Karel. *Aplikace metod detekce a rozpoznání obličeje* [online]. Brno, 2013 [cit. 2018-04-05]. Dostupné z:
https://dspace.vutbr.cz/bitstream/handle/11012/31326/xhollk00_Diplomova_prace.pdf?sequence=-1.
Diplomová práce. VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
- [10] 2D face sets. *Psychological Image Collection at Stirling* [online]. University of Stirling: University of Stirling [cit. 2018-04-26]. Dostupné z:
<http://pics.psych.stir.ac.uk/zips/Stirling.zip>

Seznam příloh

I. Obsah CD

I. Obsah CD

Součástí práce je CD, které obsahuje tyto položky:

1. Textová část bakalářské práce
2. Databáze použitých snímků
3. Algoritmus pro detekci a rozpoznání obličeje
4. Soubor popisující obsah CD

Tabulka I. 1: Popis souborů obsažených na přiloženém CD

Datový typ	Datová přípona	Název souboru
Matlab file	.m	Rozpoznani_obliceje
Matlab file	.fig	Rozpoznani_obliceje
Složka snímků		Databáze snímků
Portable Document Format	.pdf	Vojtěch_Raška_RAS0069_Bakalářská práce
Portable Document Format	.pdf	Obsah_CD

Přílohy

Laboratorní úloha

Cíl úlohy

Prostřednictvím této laboratorní úlohy se naučíte:

- Základní metody pro extrakci obličejových rysů
- Princip metody HOG

Zadání

1. Navrhněte a implementujte proceduru pro extrakci faciálních příznaků.
2. Navrhněte a implementujte proceduru pro klasifikaci faciálních příznaků.
3. Aplikujte vybrané šumové generátory na testovací obrazy.
4. Statisticky vyhodnoťte efektivitu a robustnost těchto procedur pro synteticky zašuměná data.

Předpokládané znalosti

Pro tuto úlohu se vyžaduje nastudování:

- Teoretický rozbor této úlohy

Použité vybavení

- Stolní PC
- Software Matlab s prostředím GUIDE
- Software „Rozpoznání obličeje“

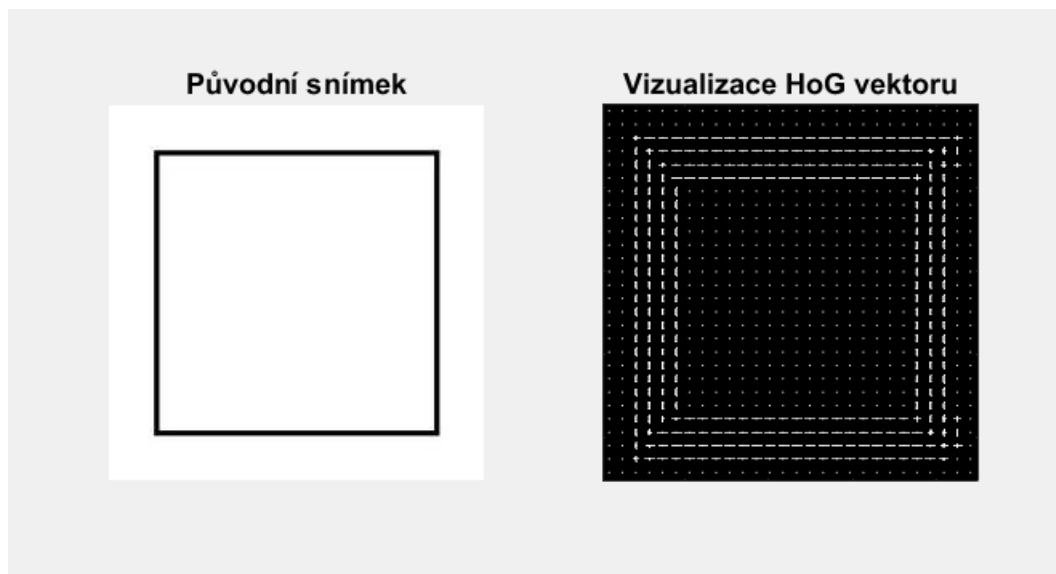
Teoretický rozbor

Biometrika je obor zabývající se identifikace subjektu na základě jeho charakteristických rysů. Dnes se využívá biometrika převážně při autentizaci. Biometrika využívá jedinečných tělesných znaků pro identifikaci osoby. Výhodou tohoto typu autentizace je, že není nutné pamatovat si několika místné kombinace hesel či neustále s sebou nosit snadno zcizitelný token, např. přihlašovací kartu. Biometrická autentizace je rychlou a pohodlnou a velice přesnou metodou, která je navíc levným řešením, vzhledem ke svým neexistujícím pozdějším nákladům. Její hlavní výhodou je skutečnost, že biometrické charakteristické znaky zůstávají během života neměnné a nelze je ukrást či zapomenout.

Metody extrakce příznaků slouží pro nalezení daných částí obličeje (oči, ústa, vrásky atp.) nalezení těch to částí pak může sloužit k identifikaci osoby nebo emocí. Obličejové příznaky lze rozdělit na trvalé a přechodové. Mezi trvalé příznaky řadíme ty části obličeje, které se na něm vyskytují trvale, tedy oči, ústa, obočí atd. Kdežto přechodové příznaky se na obličeji vyskytují pouze dočasně a vznikají při nějaké deformaci obličeje (např. úsměv).

Extrakce faciálních příznaků

Jedná se o redukci dimenzionality, která efektivně zastupuje jednotlivé body zájmu obrazu jako kompaktní vektor. Náročnost tohoto kroku se odvíjí podle velikosti databáze a rozlišení snímků. Metoda založená na extrakci HOG (Histogram of oriented gradients – Histogram orientovaných gradientů). Histogram orientovaných gradientů je deskriptor příznaků používaný pro detekci a rozpoznání obličeje. Každý pixel snímku má svou hodnotu a směr. Na základě těchto hodnot dojde k vytvoření histogramu těchto gradientů.



Obr. 1. Vizualizace HOG snímku čtverce

Pracovní postup

Postup k bodu č.1 zadání:

1. Zapněte PC a otevřete Rozpoznani_obliceje.m, program spusťte.
2. Stiskněte tlačítko „Volba databáze“ vyberte složku „Trénink“ ze složky MATLAB.
3. Pomocí tlačítka „Volba snímku“ vyberte libovolný snímek ze složky „Test“.
4. Tlačítkem „HOG“ vytvořte vizualizaci HOG vektorů, Následně uložte vizualizaci pomocí tlačítka „Uložit“ zapište si název testovaného snímku.
5. Tlačítkem „Rozpoznat“ nalezněte odpovídající shodu.
6. Kroky 3.-5. proveďte pro vždy alespoň 5 snímků pro každé natočení obličeje.
7. Z výsledných hodnot vykreslete graf míry rozpoznání pro jednotlivé natočení obličeje.

Postup k bodu č.2 zadání:

1. Pomocí tlačítka „Volba snímku“ vyberte libovolný snímek ze složky „Test“.
2. Aplikujte libovolný šum (tlačítko „Aplikovat šum“) a pomocí posuvníku nebo edit boxu nastavte parametr (pro gaušovský šum volte parametr v řádech deseti tisícín, na šum poisson nemá posuvník vliv). Šum poisson aplikujte vícekrát pomocí tlačítka „aplikovat šum“.
3. Tlačítkem „HOG“ vytvořte vizualizaci HOG vektorů, Následně uložte vizualizaci pomocí tlačítka „Uložit“ zapište si název testovaného snímku.
4. Tlačítkem „Rozpoznat“ nalezněte odpovídající shodu.
5. Kroky 1.-4. proveďte pro vždy pro alespoň 5 snímků pro 3 parametry všech tří šumů (celkově min. 45 testovaných snímků).
6. Vykreslete graf závislosti typu šumu a jeho parametru na míře rozpoznání.

Výsledky

Výsledky, grafy, komentář

Příklad vypracovaného protokolu:

**VŠB-TECHNICKÁ UNIVERZITA OSTRAVA
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

Protokol o měření

Míry rozpoznání uživatele na základě extrakce obličejových příznaků

Vypracoval: ABC0123

Datum: xx.xx.xxxx

Zadání

1. Navrhněte a implementujte proceduru pro extrakci faciálních příznaků.
2. Navrhněte a implementujte proceduru pro klasifikaci faciálních příznaků.
3. Aplikujte vybrané šumové generátory na testovací obrazy.
4. Statisticky vyhodnoťte efektivitu a robustnost těchto procedur pro synteticky zašuměná data.

Pracovní postup

Postup k bodu č.1 zadání:

1. Zapněte PC a otevřete Rozpoznani_obliceje.m, program spusťte.
2. Stiskněte tlačítko „Volba databáze“ vyberte složku „Trénink“ ze složky MATLAB.
3. Pomocí tlačítka „Volba snímku“ vyberte libovolný snímek ze složky „Test“.
4. Tlačítkem „HOG“ vytvořte vizualizaci HOG vektorů, Následně uložte vizualizaci pomocí tlačítka „Uložit“ запиšte si název testovaného snímku.
5. Tlačítkem „Rozpoznat“ nalezněte odpovídající shodu.
6. Kroky 3.-5. proveďte pro vždy alespoň 5 snímků pro každé natočení obličeje.
7. Z výsledných hodnot vykreslete graf míry rozpoznání pro jednotlivé natočení obličeje.

Postup k bodu č.2 zadání:

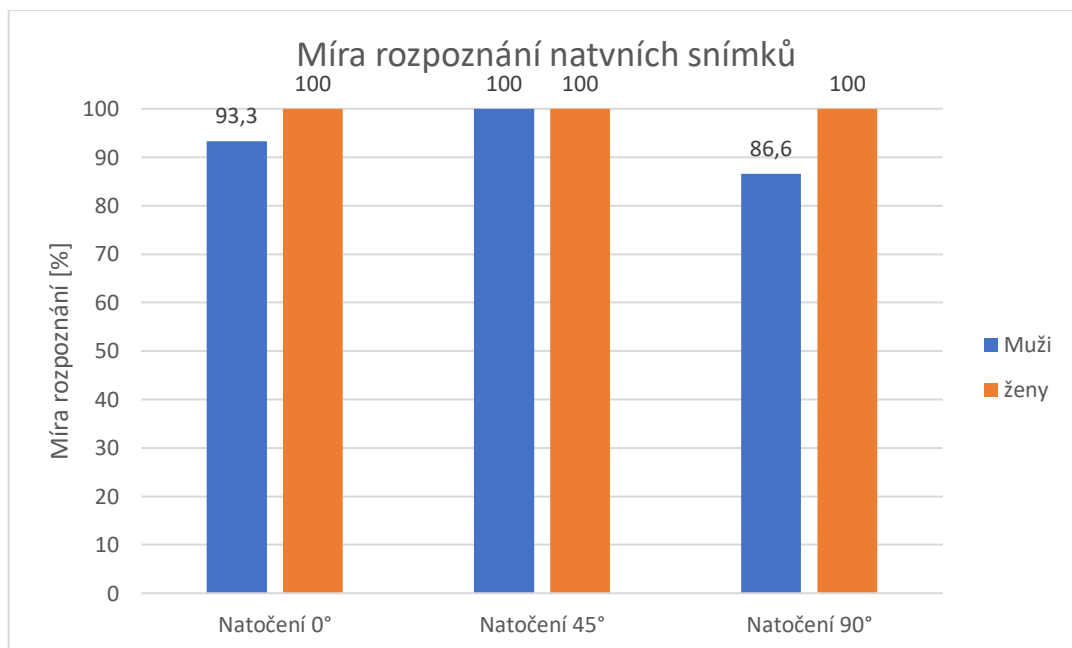
1. Pomocí tlačítka „Volba snímku“ vyberte libovolný snímek ze složky „Test“.
2. Aplikujte libovolný šum (tlačítko „Aplikovat šum“) a pomocí posuvníku nebo edit boxu nastavte parametr (pro gausovský šum volte parametr v řádech deseti tisícín, na šum poisson nemá posuvník vliv). Šum poisson aplikujte vícekrát pomocí tlačítka „aplikovat šum“.
3. Tlačítkem „HOG“ vytvořte vizualizaci HOG vektorů, Následně uložte vizualizaci pomocí tlačítka „Uložit“ запиšte si název testovaného snímku.
4. Tlačítkem „Rozpoznat“ nalezněte odpovídající shodu.
5. Kroky 1.-4. proveďte pro vždy pro alespoň 5 snímků pro 3 parametry všech tří šumů (celkově min. 45 testovaných snímků).
6. Vykreslete graf závislosti typu šumu a jeho parametru na míře rozpoznání.

Výsledky

Výsledky k bodu č. 1 zadání

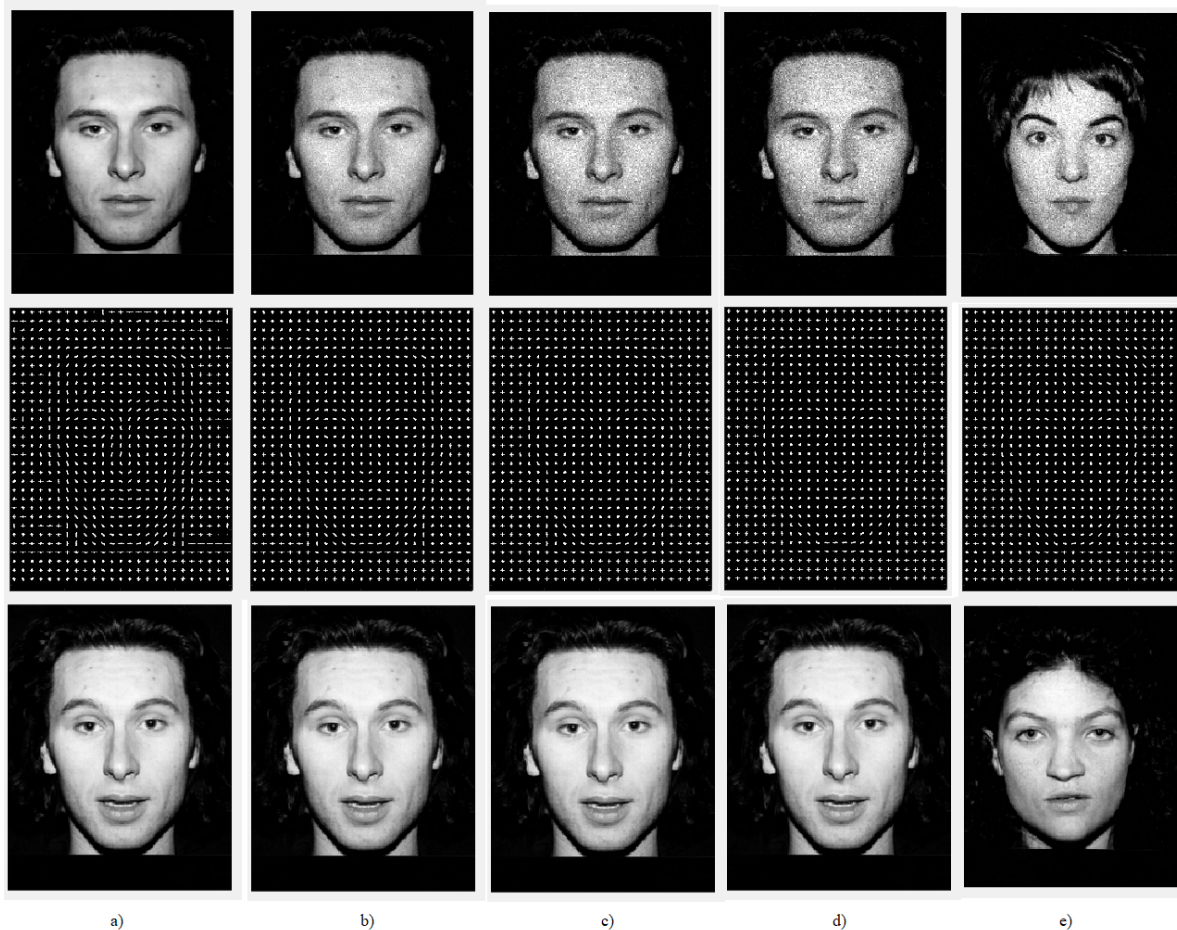


Obrázek č. 1 Příklad použitých snímků

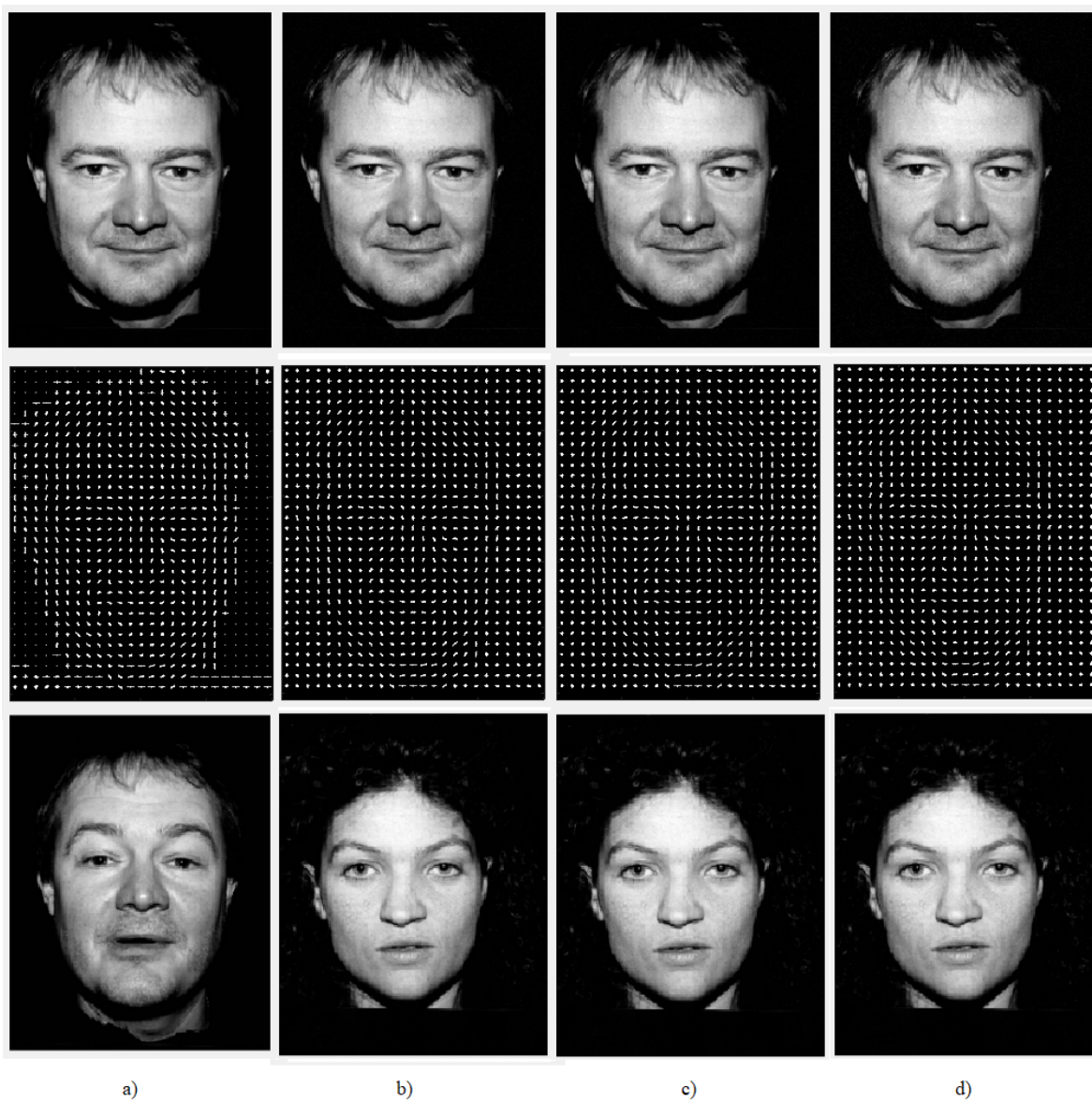


Graf č. 1 Graf výsledků na základě natočení obličeje

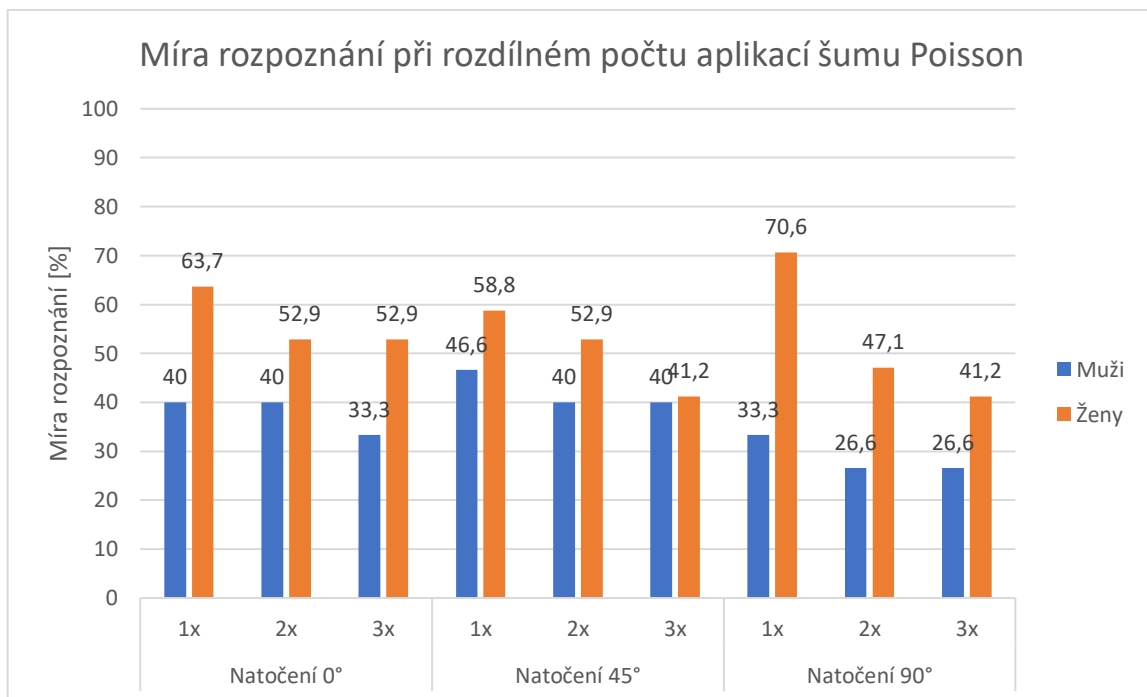
Výsledky k bodu č. 2 zadání



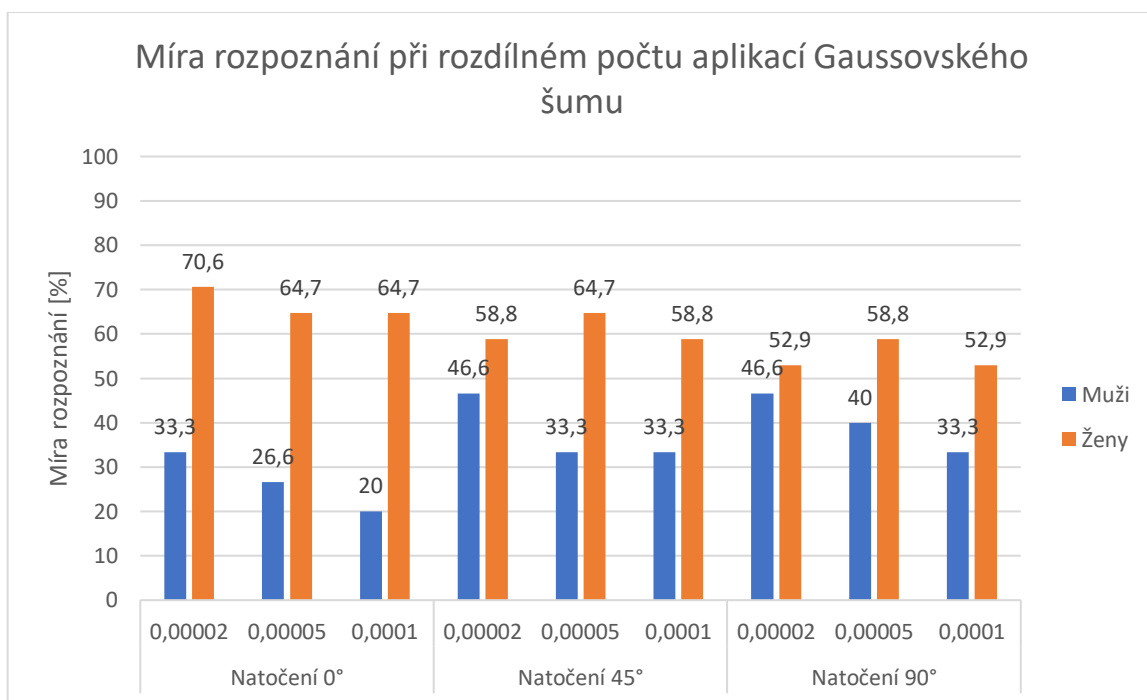
Obrázek č. 2 Příklad použitých snímků pro rozpoznání na základě aplikovaného šumu poisson



Obrázek č. 3 Příklad použitých snímků pro rozpoznání na základě aplikování gaussovského šumu



Graf č. 2 Graf míry rozpoznání na základě aplikovaného šumu Poisson



Graf č. 3 Graf míry rozpoznání na základě aplikovaného Gaussovského šumu

Závěr

Testování na nativních datech bylo provedeno pro 15 snímků s rozdílným natočením. Míra rozpoznání znázorněna na Graf č. 1 dosahuje 100 % pro všechny typy natočení pro ženy a průměrné míry 93,3 % pro muže.

Pro testování na zašuměných snímcích bylo použito Gaussovského šumu a šumu Poisson. Pro Gaussovský šum byly použity hodnoty parametru rozptylu $\sigma = 0,00002$; $\sigma = 0,00005$; $\sigma = 0,0001$. Střední hodnota je fixně nastavena $\mu = 0,0001$. Rozdíly v rozpoznání vzhledem k natočení subjektu nebyly příliš znatelné. Přesto z Graf č. 3 můžeme vidět, že i s velice nízkou hodnotou rozptylu je míra rozpoznání velice ovlivněna. Míra rozpoznání pro ženy se pohybovala v rozmezí 70-52,9 % a můžeme vidět, že míra rozpoznání se postupně snižovala s rostoucí hodnotou rozptylu. Míra rozpoznání pro muže byla znatelně nižší a pohybovala se v rozptylu 46,5-20 %.

Při aplikaci šumu Poisson byla míra rozpoznání opět největší pro ženy a to 70,6 % pro muže byla nejvyšší míra rozpoznání 46,6 % tento rozdíl může být dát větší variací účesů u žen a výraznějším rozdílům v rysech obličeje. Opět lze vidět, že míra rozpoznání se s rostoucím počtem aplikací šumu pozvolna snižuje. Výsledky aplikace šumu Poisson lze vidět na Graf č. 2.